华章数学译丛



60

ELSEVIER
爱思唯尔

An Introduction to the Mathematics of Financial Derivatives

Third Edition

金融衍生工具数学导论

（原书第3版）

[美] 艾利·赫萨（Ali Hirsa） 萨利赫 N. 内夫特奇（Salih N. Neftci） 著

冉启康 葛泓杉 李君格 译



华章数学译丛

60

An Introduction to the Mathematics of Financial Derivatives

Third Edition

金融衍生工具数学导论

（原书第3版）

[美] 艾利·赫萨（Ali Hirsa） 萨利赫 N. 内夫特奇（Salih N. Neftci） 著

冉启康 葛泓杉 李君格 译

机械工业出版社
China Machine Press

图书在版编目（CIP）数据

金融衍生工具数学导论（原书第3版）/（美）艾利·赫萨（Ali Hirsa），（美）萨利赫 N. 内夫特奇（Salih N. Neftci）著；冉启康，葛泓杉，李君格译．—北京：机械工业出版社，2016.7

（华章数学译丛）

书名原文：An Introduction to the Mathematics of Financial Derivatives, Third Edition

ISBN 978-7-111-54460-9

I. 金… II. ①艾… ②萨… ③冉… ④葛… ⑤李… III. 金融衍生工具－经济数学－研究 IV. F830.9

中国版本图书馆CIP数据核字（2016）第182401号

本书版权登记号：图字：01-2015-0847

An Introduction to the Mathematics of Financial Derivatives, Third Edition
Ali Hirsa and Salih N. Neftci
ISBN: 978-0-12-384682-2

本书是一本优秀的金融衍生工具方面的教材，对现代资产定价理论所需的基本数学工具进行了系统全面的介绍，包括：套利定理、风险中性概率、用于金融领域的微积分、鞅、偏微分方程、Girsanov定理和Feyman-Kac公式等．本书案例丰富、体系精巧，对资产定价相关的数学原理和方法的描述简明直观．适合作为高等院校财经类、数学类专业的本科或者研究生教材，也适合从事金融工作的在职人员阅读．

出版发行：机械工业出版社（北京市西城区百万庄大街22号 邮政编码：100037）

责任编辑：和　静	责任校对：董纪丽
印　　刷：北京诚信伟业印刷有限公司	版　　次：2016年8月第1版第1次印刷
开　　本：186mm×240mm　1/16	印　　张：28.5
书　　号：ISBN 978-7-111-54460-9	定　　价：99.00元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换
客服热线：(010) 88378991　88361066　　投稿热线：(010) 88379604
购书热线：(010) 68326294　88379649　68995259　　读者信箱：hzjsj@hzbook.com

译 者 序

美国金融数学专家 Ali Hirsa 与 Salih N. Neftci 合著的《An Introduction to the Mathematics of Financial Derivatives》是一部在美国和欧洲非常流行的著作．本书使用随机微积分知识讨论金融衍生品定价理论，简单易懂、内容丰富．作者以高超的手法对金融衍生品定价中所需的数学理论进行了全面的描述和总结，并系统地介绍了金融衍生品的常用定价方法和最新进展．两位作者都有多年业界工作经验，与此同时又都热爱教学．Ali Hirsa 为哥伦比亚大学与纽约大学兼职教授，Salih N. Neftci 生前先后工作于纽约城市大学(CUNY)、新学院(The New School)、Baruch College、国际货币基金组织、世界银行等．

近几年来，金融衍生品定价理论在国际金融界和数学界受到了越来越广泛的重视，国内外出版了大量有关金融衍生品定价理论方面的专著和教科书，然而这些书中，在阐述其主要内容(如关于期权的定价理论等)时，大都直接或间接地使用了随机过程、随机分析、高级计量经济学、运筹学等现代数学知识，并且把这些知识作为读者已经掌握的东西．而另一方面，目前一般大学本科生所掌握的数学工具主要是微积分、线性代数和初等概率论，此类著作中涉及的现代数学知识远远超出了包括数学专业在内的大学生的知识范畴，甚至在金融投资部门从事实际工作的专业人员也难以适从，本书以其起点低、直观易懂的特点使人感到眼前一亮．它系统、全面地介绍了金融衍生品定价理论的基本内容．并将读者应该具备的数学基础严格限定在包括经济、金融、管理等专业的绝大多数本科生的水平．由于作者在内容选择、结构安排和逻辑体系设计方面的精巧构思，所以能以相对较少的篇幅，把书中所讨论的问题的经济背景以及解决这些问题的数学方法和基本思想，系统而又简洁明快地展示给读者，其中某些问题的讲述还具有相当的深度．相信那些从事实际工作的读者以及对该学科感兴趣的在校本科生或研究生读者会大为受益，本书适合作为高等院校财经类专业、数学类专业以及学习过微积分、概率论课程的其他专业的本科学生或研究生的教材，同时也适合从事金融工作的在职人员阅读．

受机械工业出版社华章分社之托，我们将此书的第 3 版译成中文．全书由上海财经大学数学学院冉启康教授，硕士研究生葛泓杉、李君格共同翻译，王佳捷老师参与了校对工作。在翻译过程中，我们得到机械工业出版社华章分社王春华编辑的大力帮助，在此表示衷心的感谢！限于时间和水平，译文的不当之处在所难免，敬请本书的读者和有关领域的专家批评指正．

译者
2016 年 2 月

符号和缩写列表

R	实数集
$\mathbf{R}^{+}$	正实数集
Z	整数集
N	自然数集
$p(t, T)$	T 时刻到期的零息债券在 t 时刻的价格
r_t	日历时间 t 的瞬时短期利率
B_t	0 时刻以 1 美元为本金，按瞬时短期利率计息，到时间 t 所得的本息之和
$f(t, T)$	日历时间 t 所观察到的未来$[T, T+\mathrm{d}t]$时段内的瞬时远期利率
P	现实世界(真实)测度
Q	风险中性测度
$E(x)$	x 在某测度下的数学期望
$E_t(x)$	已知截止到时间 t 的所有信息，x 在某测度下的条件期望

目 录

第 1 章　金融衍生品概论

1.1 引言

本书是一本介绍金融衍生品量化方法的读物，它是一本比较简单而且具有启发性的从数学概念入手，将数学应用于金融市场的读物．

因为这是一本讨论资产定价背后的理论依据的书，在阅读本书的诸多案例时，读者需要理解资产价格的严谨定价方法，所以我们有必要对所研究的证券作简要的介绍，本章就是为了这一目的而编写的．读者可以阅读一些其他的书籍来获得更多有关衍生品的背景知识，Hull(2009)是一本很好的参考读物，Jarrow 和 Turnbull(1996)介绍了另外一种方法．Ingersoll(1987)和 Duffie(1996)的内容更深入了一步，它们介绍了衍生品与其价格之间的内在理论联系．Das(1994)是一本关于衍生品合约的实用内容指南．此外，Wilmott(1998)介绍了有关衍生品定价的一些新方法．

本章首先介绍金融衍生品的两个基本概念：期权和远期(期货)；接下来，我们将介绍更复杂的衍生品：互换交易；最后，我们将证明：一个复杂的互换交易可以分解为若干个远期和期权的组合．这种分解是非常有用的，因为只要我们能够对远期和期权定价，我们就能重组出任何互换交易并定价．本章还将给出一些记号，这些记号将贯穿全书．

1.2 定义

用从业者的话来说，"衍生证券是一种金融合约，其价值是根据其他一些现金市场工具(如股票、债券、现金、货物)的表现衍生出来的"．[⊖]

学术上对衍生品的定义更加严谨．

定义 1 (Ingersoll，1987)．*对于一份金融合约，如果它在到期日 T 的价格完全取决于 T 时刻某标的现金工具的市场价格，那么就称该合约是一种衍生证券，也叫未定权益．*

根据定义，在衍生合约的到期日 T，衍生资产的价格 $F(T)$ 完全取决于标的资产的价格 S_T，在到期日之后，该证券不复存在．衍生品的这一简单性质在定价过程中起到了非常重要的作用．

在本书中，我们用 $F(t)$ 或 $F(S_t,t)$ 表示衍生品的价格，S_t 表示标的资产的价格，t 表示时间．金融衍生品有时候会获得收益 dt，有时候没有收益，T 总表示到期日．

1.3 衍生品的分类

我们可以将衍生证券分为三大类：

⊖ 参见 Klein 和 Lederman(1994)，第 2～3 页．

1）期货和远期合约．

2）期权．

3）互换．

远期合约和期权被视为两大基本衍生证券，互换和其他一些复杂证券被称为混合证券，最终能被分解为一些基本的远期合约和期权．

我们用 S_t 表示相关金融工具（即标的资产）的价格．

主要有以下5种标的资产：

1）股票：是生产商品或提供服务的企业所发出的具有真实回报的所有权凭证．

2）货币：是指来自政府或银行的债务，它对实体资产没有直接的所有权．

3）利率：实际上，利率并非资产，因此需要设计一种名义资产，以便人们应对将来的利率变化，欧洲美元期货就是一个例子．此类目中还包括债券、票据、短期国债的衍生品．债券、票据和短期国债是政府的借债工具，政府承诺在规定日期对所有人支付相应金额．通过操作债券、票据、短期国债的衍生品，人们可以应对将来各种利率的变化．在大多数情况下㊀，这类衍生工具是会生效并产生标的资产的实际交付，而不是仅停留在纸面上．

2

4）指数：S&P 500股价指数和FTSE100股价指数是两个股票指数的例子，CRB商品指数是一种商品价格指数．这些指数本身不是资产，但指数合约具有名义金额并且能够被买卖，其价格变化对应于标的指数的变化．

5）大宗商品：主要的类目包括

- 软商品：可可、咖啡、糖．
- 谷物及油籽：大麦、玉米、棉花、燕麦、棕榈油、马铃薯、大豆、冬小麦、春小麦等．
- 金属：铜、镍、锡等．
- 贵金属：黄金、铂金、银．
- 家畜：牛、猪、五花肉等．
- 能源：原油、燃油等．

这些标的商品并非金融资产，它们在性质上属于商品．因此在大多数情况下，这些商品会参与实际交易和贮藏．

此外，还有另外一种分类标的资产的方法，该方法对我们学习本书很有帮助．

1.3.1 现金交易市场

有些金融衍生品的标的资产属于现金交易市场，这类标的资产包括黄金、白银、货币、长期国债等．

在这类市场中，人们可以以无风险的利率借入资金（需要实物资产抵押）来购买并储藏现金交易市场中的商品，并安全持有直至衍生合约到期．另一方面，人们也可以轻松地获

㊀ 在巴黎，有大量“名义”法国政府债券的交易．

得这些商品的远期合约或期货合约.

例如，人们可以以无风险利率借入资金来购买长期国债，并且在相应的期货合约的交割日前持有该国债.这就等价于购买一份期货合约并在到期日进行交割.人们也可以对货币、黄金、白银、原油等商品作类似的操作.⊖

纯粹的现金交易市场还有一个性质，关于标的工具未来需求与供给的信息，不影响现价和远期价格的价差.这种价差只依赖无风险利率的水平、贮藏和保险费用等，标的工具供需情况的信息对现价和远期价格造成等额的变化.

1.3.2 价格发现市场

第二种标的资产来自价格发现市场.在这里，标的工具在期权到期日之前不会产生实物交易或贮藏，这类货物可能是易于腐坏不宜贮藏的，也可能是在衍生品交易时根本不存在.一个典型的例子是春小麦，当春小麦的期货合约签订时，对应的现金市场还不存在.

借入、购买、贮藏资产直到未来的某个到期日从而盈利，这样的策略在价格发现市场中不可行.在这些条件下，任何关于未来标的商品供给和需求的信息不会影响对应的现金价格.这种信息会在未来市场中被人们发现，包括合约期间.

1.3.3 到期日

衍生品价格 $F(t)$ 和标的资产价格 $S(t)$ 的关系仅在到期日 T 明确已知.对于远期合约或期货，我们自然地假定

3

$$F(T) = S_T \tag{1.1}$$

也就是说，在到期日 T，期货合约的价值应该等于其现金等价物.

例如，一份承诺交付 100 盎司黄金的(场内交易)期货合约，在到期日其价值应该和现金交易市场上 100 盎司黄金的价格保持一致.在时间 T，它们完全等价，因此，对于黄金期货，我们可以确切地说等式(1.1)是成立的.

当 $t<T$ 时，$F(t)$ 可能不等于 S_t，但我们可以建立一个函数，得到 S_t 和 $F(t)$ 的关系.

1.4 远期合约和期货

期货和远期合约是线性工具，下面我们将首先讨论远期合约，最后再简要介绍期货和远期合约的区别.

1.4.1 远期合约

定义 2 远期合约是一个在确定的将来时刻按确定的价格购买或出售某项资产的协议.

⊖ 然而，对原油来说，储藏费用非常高，这一切是环境等因素造成的.

双方的合约必须指明交割日和远期价格. 如果远期购买成立，该合约的购买方称作标的资产的多头方. 如果在交割日现价比远期价格更高，多头方盈利，否则，蒙受损失.

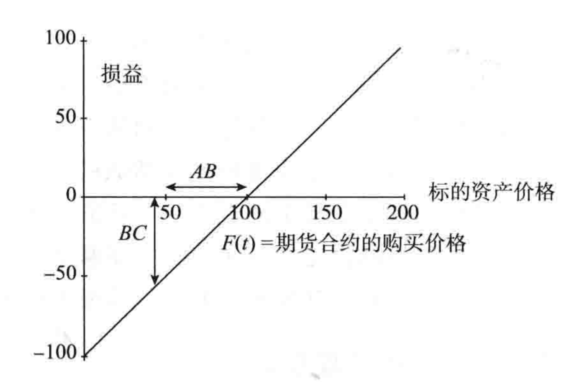


图 1-1 远期合约中多头方的收益图

图 1-1 简单给出了多头方损益与标的物现价的关系. 合约在 t 时刻以 $F(t)$ 价格被购买，假定合约 $t+1$ 时刻过期，向上倾斜的直线显示交割日多头方的损益. 直线斜率为 1.

4

如果 S_{t+1} 大于 $F(t)$，那么多头方将获利[⊖]. 如果直线斜率为 1，线段 AB 的长度等于垂线段 BC，$t+1$ 时刻的损益可以用 BC 线段的长度来表示.

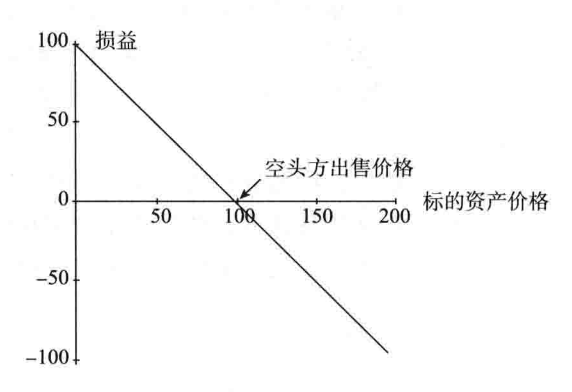


图 1-2 远期合约中空头方的收益图

图 1-2 给出了同样条件下空头方的损益情况.

这种损益图对理解衍生品原理很有帮助，本书不作进一步介绍，读者可以阅读 Hull(2009)的书以便进一步了解这些内容.

1.4.2 期货

期货和远期合约很类似，主要有以下区别:

期货通常在交易所内交易，交易所详细规定期货合约的标准化条款，设定若干具体交割日期. 远期合约交易事项由双方协商确定，在场外进行交易.

交易所负责结算期货交易，并提供复杂的机制以减少违约风险.

此外，期货合约实行逐日盯市制度. 也就是说，每天交易结束时，投资者的保证金账户要进行调整，以反映其盈利或损失.

1.4.3 回购协议、反向回购协议及弹性回购协议

附买回协议，又称回购协议，是指证券交易的双方在进行交易的同时，以契约的方式约定在将来的某一日以约定的价格，由证券的卖方(正回购方)向买方(逆回购方)购回该笔证券的交易行为，回购价格和原价格的差异取决于回购利率. 证券的购买者相当于债权人，而出售者相当于债务人，证券就相当于这笔有固定利率的债务的抵押. 一份回购协议可以视为现货销售和远期合约的组合. 在现货销售中贷款人向借款人提供货币，债券则被合法

⊖ 由于合约在 $t+1$ 时刻到期，S_{t+1} 将等于 $F(t+1)$.

转让给贷款人作为交换．远期合约保证了借款人付清贷款以及贷款人退还债券．远期合约价格和现货价格之差相当于借款的利息，而远期合约的交割日是这笔债务的还款日．

回购协议可分为三种：(a)隔夜回购，(b)定期回购，(c)不定期回购．隔夜回购的到期日为第二天，定期回购的到期日为一个特定的时间，不定期回购没有明确的终止日期．回购交易通常有以下三种形式：递送、第三方协议回购、持有保管(这种情况下出售方在回购期间持有证券)．持有保管主要用于防范以下风险：逆回购方在协议到期前破产，则正回购方无法获得作为抵押物的证券．

出售证券并承诺将来回购的一方，成为*正回购方*．购买证券并承诺将来出售的一方，成为*逆回购方*．

回购通常用来筹集短期资金，属于货币市场工具．

弹性回购协议是一种回购时间灵活机动的回购协议．该协议下，逆回购方可以在一定范围内提前将证券卖给回购方，而非恰好在约定日的当天，提前执行的金额和时间都是灵活的．

弹性回购和传统回购主要有 4 个区别：

- 由于现金可提取，具有凸性．
- 像交易一样正式签订条款．
- 更多的合同(有额外的合同提供信用支持和破产保护)．
- 在弹性回购市场，合同的一方是市政债券发行者．

5

弹性回购主要分为两类：有担保的，没有担保的．

对于担保弹性回购，逆回购方将收到抵押物，其委托人将为其监管并结算抵押物，抵押物一般是国库券、GNMA 债券、机构 MBS/CMO 等．用作抵押物的证券种类在相关文件和求证建议书(RFP)中都有具体的介绍，在大多数情况下，抵押物来自逆回购市场，例如，银行将一笔证券买来并签订协议将其出售回去．

对于非担保弹性回购，客户并不会收到抵押物，在这种情况下，客户可能会得到更高的回报．一笔非担保弹性回购的平均面额一般在一到二千万美元，其金额一般低于担保弹性回购．

1.5 期权

期权是资产定价中的第二个基本衍生品．在后面的章节中，我们经常用定价模型去研究看涨期权的价格，以此作为介绍随机积分的例子．

远期合约和期货规定了合约双方在交割日对标的资产进行交割的义务．而期权所有者拥有执行交割的权利，而非义务．

期权主要分为以下两类：

定义 3 *某证券 S_t 的欧式看涨期权是指：在到期日 T 以执行价格 K 购买一份该证券的权利，该权利只有在到期日才能实施，且可放弃．在到期日之前的时刻 t，该期权可以以价*

6

格 C_t 出售出去，C_t 又称权利金．

欧式看跌期权与之类似，只是所有者的权利改为在到期日以执行价格出售标的资产．

与欧式期权不同的是，美式期权能在签署日和到期日之间的任何时刻被执行．

基于多种原因，交易者和投资者需要算出看涨期权的无套利价格．欧式看涨期权的价格 C_t 在期权签订之前是未知的，交易者希望计算出期权的发行价格．假如期权是一种可交易的证券，它将在交易中形成市场价格，假如期权交易在场外进行，它也将被频繁交易，价格也是易于观察的．

然而，期权交易也许并不频繁，有时为了评估风险，交易者可能想要知道期权每天的价格 C_t，有时交易者认为市场低估了某种期权，而这种低估可以用于获利，在这些时候，我们就需要知道期权的无套利价格 C_t.

一些记号

要得到看涨期权的价格 C_t，最理想的方法是找到 C_t 关于标的资产价格以及其他相关变量的封闭式公式．

在 t 时刻，关于 C_t，我们已知的仅有它在交割日 T 的值．事实上，

- 如果不存在交易手续费，并且
- S_t 和 C_t 的买卖差价是 0，

那么在到期日，C_t 的值只可能有两种情况：

如果期权在到期日处于虚值状态，也就是说，

$$S_T < K \tag{1.2}$$

那么期权的价值为 0. 人们能够以 S_T 从市场上购买标的资产，而这小于执行价格 K. 期权持有者不会执行该期权，即不会以价格 K 购买标的资产．因此，

$$S_T < K \Rightarrow C_T = 0 \tag{1.3}$$

但如果该期权处于实值状态，也就是说，在时刻 T,

$$S_T > K \tag{1.4}$$

那么该期权有价值．所有者应该执行该期权，他可以以价格 K 买入标的资产再以更高的价格 S_T 卖出．由于没有交易费用和买卖差价，净利润为 $S_T - K$. 了解该规则的市场参与者就会将该期权定价为 $S_T - K$，因此，

$$S_T > K \Rightarrow C_T = S_T - K \tag{1.5}$$

上述两种情况可以概括为

$$C_T = \max[S_T - K, 0] \tag{1.6}$$

它表示 C_T 取值为括号内两个值中较大的那

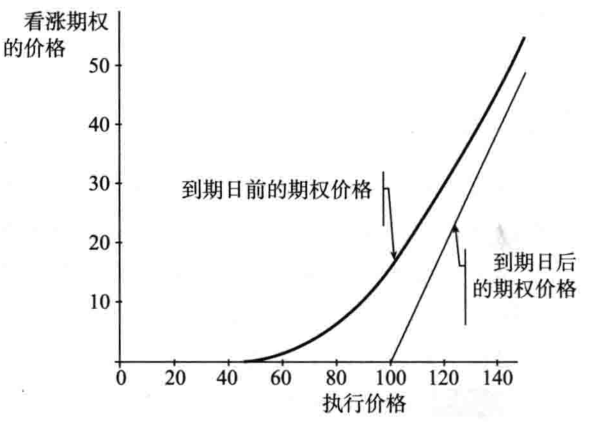


图 1-3 到期日前和到期日股票价格与期权价格的关系

一个．在后面的章节中，我们将经常使用该记号．

图 1-3 标出了(1.6)式中 C_T 与 S_T 的关系：当 $S_T \leqslant K$ 时，C_T 为 0；当 $K < S_T$ 时，C_T 与 S_T 以相同速率增长，因此当 $K < S_T$ 时，C_T 与 S_T 的关系是一条斜率为 1 的直线．期权是非线性工具．

图 1-4 标出了到期日前不同时刻看涨期权的价格图形．当 $t < T$ 时，期权价格函数是一条光滑连续的曲线．而在到期日，期权的价格函数是以执行价格 K 为角点的分段线性函数.

7

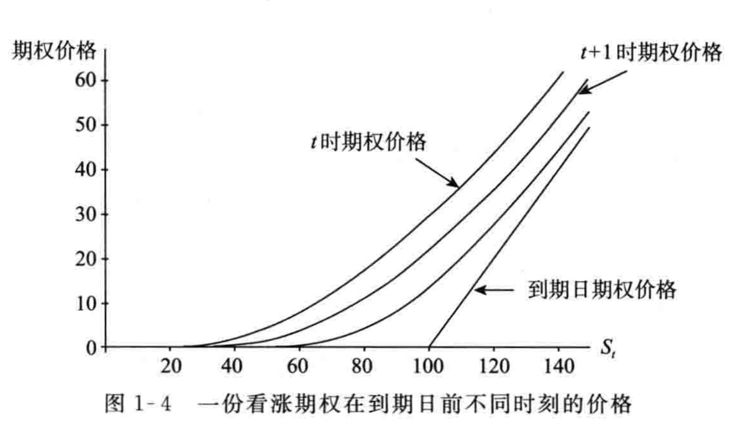


图 1-4　一份看涨期权在到期日前不同时刻的价格

1.6　互换

互换和互换期权是常见的衍生品，但这并不是我们感兴趣的原因．一个常用的对互换和互换期权定价的方法是：将其分解为一系列远期合约和期权．该方法体现了远期合约和期权作为基本衍生品的特殊功能，也解释了为什么我们前面的章节把重点放在了这两个问题上．

8

定义 4 互换是指同时出售和购买现金流，该现金流可以是不同的币种、不同的利率，也可以是一些其他金融资产．

对一般互换工具的介绍超出了本书的范畴．正如前文所述，我们的目的是介绍衍生品定价所包含的数学模型，而非衍生品本身．所以，我们仅重点介绍一个有代表性的例子．

1.6.1　一个简单的利率互换

将一份互换分解为若干部分是金融工程学和资产定价学的一个典型案例．它还体现了远期、期权的特殊地位．我们将仔细讨论利率互换．读者可以从 Das(1994)的书中查阅到更多高级的互换结构．

现在我们讨论一个最简单的利率互换．交易双方 A、B 将进行如下步骤：

1) 甲方 A 需要 100 万浮动利率的贷款，乙方 B 需要 100 万固定利率的贷款，但由于市场条件以及种种原因，B 在获得浮动利率方面比较具有优势．

2）A和B决定利用这一比较优势，双方各自以具有比较优势的利率从市场上获取贷款，并且支付对方的利息．

3）甲方A以固定利率借得100万元，A将向B收取这笔借款的利息，然后付给银行．

4）乙方B将以浮动利率借得100万元，B将向A收取这笔借款的利息，然后付给银行.

5）请注意双方的贷款本金是相同的，都是100万元．因此，不需要交换这100万元．这笔钱叫做名义本金．又因为利息也是用相同的币种来支付，所以交易双方只需要交换利息差额．这就是利率互换．

这个基础的利率互换实际上是利息支付的交换．双方各自通过擅长的方式借款然后交换支付利息．最终双方都能获得更低的利率，中间人也能赚取佣金．

互换总可以分解为一系列更简单的远期合约，这些合约重新组合，就是互换．我们可以分别对这些远期合约定价，进而对互换定价．这种将互换分解为远期的方法极大地推动了互换合约定价的发展．

1.6.2 可取消互换

可取消互换是一种双方均有权利在到期日前单方或双方协商取消互换的互换协议，可取消的日期将在合同中提前规定．有以下两种可取消互换：可要求互换，可提出互换．

可要求互换是指在固定利率与浮动利率互换中，支付固定利率的一方有权在合同规定的可取消日期中取消互换，并且不用支付代偿还金额的现值，就可以取回保证金．在可提出互换中，相反，支付浮动利率的一方有权在一定期限内取消互换．普通互换的交易方如果要取消互换，一般必须要支付未支付金额的净现值．

有些债务可以在到期日之前提前偿还本金，如可取消债券．拥有这一类债券的机构往9往会选择可取消互换．可取消互换通常用于对冲，公司可以利用提前支付权和可取消互换来避免他们的资产、具有提前还款权的债务以及互换不匹配的情况．

1.7 小结

本章我们回顾了一些基本的衍生工具．我们有这样两个目的：第一，介绍基础的衍生证券以便后面的举例；第二，引入衍生资产定价中的一些记号．在后面，我们往往是先得到远期合约、期权等基础衍生品的定价公式，再把结构更复杂的衍生品分解为一系列远期合约和期权，通过这种方法，基础衍生品的定价能够对复杂衍生品的定价起到直接的帮助作用．

1.8 参考阅读

Hull(2009)非常值得借鉴，它在很多方面是独一无二的，实践者可以把它当作工具书来查阅，初学者可以把它当作教材来学习，它易于理解，论述严谨．Jarrow 和 Turbull (1996)是一本深受欢迎的提高性读物．Duffie(1996)详细介绍了动态资产的定价理论，但

它在实际应用方面的介绍并不详细，不过数学功底扎实的操作者也许会觉得它很有用．Das(1994) 给出了很多衍生品工具的案例．

1.9 习题

1. 考虑以下金融工具：

- 投资者 A 以价格 S 卖出一只股票，同时买入一份该股票的执行价格为 K 的欧式看涨期权．
- 投资者 B 购买一份执行价格为 K_1 的欧式看跌期权和一份执行价格为 K_2 的看涨期权，$K_1 \leqslant K_2$.
- 投资者 C 买入一份欧式看跌期权同时卖出一份看涨期权，执行价格均为 K_1；并且买入一份看涨期权同时卖出一份看跌期权，执行价格均为 K_2，$K_1 \leqslant K_2$.

(a) 画出每种情况下到期日的收益图．

(b) 这些图在到期日之前是什么样子的？（参见图 1-3 和图 1-4.）

2. 考虑一个支付固定利率的普通利率互换交易方，已知：

- 该互换 12 个月后开始，24 个月后结束．
- 浮动利率是 6 个月美元 Libor 利率．
- 互换利率为 $K=5\%$.

(a) 画出本次互换的现金流向图．

(b) 写出包含两个远期利率合约(FRA)的组合合约，使之与上述互换等价，并详细介绍其中的参数．

(c) 你能用合适的期权组成一个互换吗？

3. 设 1 吨小麦的无套利 3 个月期货价格为 F_t，假设 1 吨小麦 1 年的贮藏费用是 c 美元，管理费用是 s 美元，对交易者而言现货小麦的可接受利率是 $r\%$，最后，假设小麦没有便利收益．

(a) 写出 F_t 的公式．

(b) 令 $F_t=1500$，$r=5\%$，$s=100$ 美元，$c=150$ 美元，小麦的现货价格为 $S_t=1470$ 美元，F_t 是无套利的吗？构造一个有套利的方案．

(c) 假设所有参数保持不变，该方案到期日的收益是多少？ 10

4. 一只股票现价为 $S_t=100$，平价欧式看涨期权的价格为 3，平价看跌期权价格为 3.5. 假设没有交易费用且股息为 0，交易者可以以 5%的年利率借入或借出资金，市场是流动的．

(a) 交易者签订了一份该股票的远期合约，12 个月后交割，该远期合约的价格是 F_t，那么 F_t 是多少？

(b) 假设市场对该远期合约的估价是 $F_t=101$，试建立 2 个有套利的资产组合．

5. 如果投资者对资产的未来状况持以下观点，那么他们会分别采取怎样的欧式看涨、看跌期权组合加以应对？

(a) 未来较高的波动率．

(b) 未来较低的波动率．

(c) 可能发生偶然事件，对资产价格产生消极影响．

对于每种资产组合，画出相应的损益图．

6. 针对某股票在期权到期日不同可能的股价情况，画出牛市看涨差价期权损益函数的图像． 11

第2章 套利定理入门

2.1 引言

目前对于衍生资产的所有定价方法都使用了套利这一概念，在套利定价方法中更是直接用到了套利概念．资产的价格是在排除了套利机会的条件下得到的．在均衡定价方法中，无套利假定是一般均衡条件之一．

简单来讲，套利就是通过对不同资产持有不同仓位，来取得高于美国联邦短期国库券收益的无风险收益的行为．如果这种无风险收益可以被实现，就称为存在套利机会．

套利机会通常以两种形式存在．第一种形式为，我们可以在不获得当期收益的情况下进行一系列投资，以期在之后获得正收益．例如，卖空一只股票，然后用所得收益购买以该股票为标的资产的看涨期权．在该投资组合中，分别建立了一个看涨期权的多仓和一个基础股票的空仓．如果操作得当，空仓与多仓的未来资产总值可以是确定的，从而投资组合不受风险影响．在佣金和交易费被减去后，理论上这种投资将不会产生额外利润．否则，我们就称第一类套利机会存在．

在第二种套利机会中，投资组合在当期可能有负收益，但在将来会得到非负收益．

通过上面的概念，我们将能引入一个非常实用的定义——金融资产的“公平价格”．对于一种证券，如果它的定价不会导致两类套利机会的存在，那么就称该定价是公平的．这种无套利机会的资产定价将会被作为基准，一旦资产价格发生偏离，就意味着可能获得额外收益．

在实际中，套利机会是可能存在的，但这并不影响我们研究无套利价格的兴趣．事实上，确定无套利价格是研究衍生品价值的核心．我们可以想到至少4种无套利价格的用途.

首先，如果衍生品开发机构打算开发一种新型金融产品，由于这种产品此前没有交易记录，它的交易价格无法通过观察金融市场的实际交易情况得出．这时，可以利用无套利价格对该产品定价．

其次，无套利价格可以应用在风险管理中. 在通常情况下，风控经理们会通过模拟发生“最坏的情况”来衡量投资组合的风险，并且这种模拟过程会定期进行．考虑到在模拟过程中所出现的问题都是对未发生事件的假设，每次都会需要使用某基准价格，这时就会用到无套利价格．[⊖]

再次，在对投资组合中的资产进行逐日盯市时，由于财务主管需要获得一些可能最近都未发生交易的低流动性资产的当前市场价格．这时，计算其相应的无套利价格可以作为

[⊖] 注意，这些设计出来的情境不一定都是正确的. 例如，我们并不能确定当市场遭遇极端事件的冲击时，无套利条件所要求的流动性是否能够满足．

一种解决办法．

最后，无套利定价计算的基准价格可以与实际交易中的交易价格进行比较．两者间的显著差异可能预示着存在获得额外收益的机会．通过这种方法，无套利价格可以用来捕捉极短时间内出现的不合理交易价格，若无套利价格高于交易价格，则该衍生品价格就被低估了，应该建仓．反之，则该衍生品价格被高估了．

在无套利理论基础上建立起来的一系列数学模型是计算这种基准价格的主要工具．

2.2 记号

本节我们将通过引入一些记号对讨论的问题符号化，记号是任何一种完整的数学方法都必不可少的．正确理解记号有时和理解定理所蕴含的数学逻辑同样重要．

14

2.2.1 资产价格

符号 t 表示时间．期权、期货、远期和股票等证券的资产价格将由一列向量表示，记为 $\boldsymbol{S}_t$. 金融市场的任何证券都可以放在该数组中：

$$S_t = \begin{bmatrix} S_1(t) \\ \vdots \\ S_N(t) \end{bmatrix} \tag{2.1}$$

其中，$S_1(t)$可以表示无风险借贷资产的价格，$S_2(t)$可以表示一种特定股票的价格，$S_3(t)$可以表示以该股票为标的资产的看涨期权的价格，$S_4(t)$可以表示相应的看跌期权的价格，等等. S_t 中的下标 t 表示该价格为 t 时刻的价格．在离散时间条件下，证券价格可以表示为 S_0，S_1，…，S_t，S_{t+1}，…. 然而，在连续时间条件下，下标 t 可以为 0 到无穷大之间的任意值．通常我们写成

$$t \in [0,\infty) \tag{2.2}$$

一般情况下，0 代表初始时刻，t 代表现在．如果有

$$t < s \tag{2.3}$$

那么 s 表示将来某一时刻．

2.2.2 状态

为了余下章节讨论的需要，我们引入状态的概念．刚开始接触时，这个概念可能有些抽象，但它与实际是紧密联系的．

我们用向量 W 表示所有可能状态，

$$W = \begin{bmatrix} \omega_1 \\ \vdots \\ \omega_N \end{bmatrix} \tag{2.4}$$

其中ω_i 表示每一个可能发生的结果．这些状态都是相互独立的，并且其中至少有一种情况

会发生.

一般情况下，金融资产在不同状态 ω_i 下会有不同的价值，并产生不同的收益．我们假定金融市场共有 K 种可能的状态(K 为自然数).

很容易将这一概念形象化．对于交易员来说，可能获得收益的唯一时刻就是在“下一刻”. 显然，证券价格是会发生变化的，而我们并不知道它的走向．但是，在一个小的时间间隔内，证券价格可能会上涨，也可能会下跌，还有可能不发生任何变化．因此，我们可以认为共有三种可能的状态．

2.2.3 收益和回报

由于不同状态下证券的收益是不同的，因此，状态 ω_i 很重要．我们用 d_{ij} 表示每单位的证券 i 在状态 j 下的收益，它由两部分组成．

第一个部分是资本损益．资产价值可能发生升值或贬值．对于“看多”某种资产的投资者来说，升值会产生资本收益，而贬值会造成资本损失．对于“看空”某种资产的投资者来说，资本损益正好相反．[⊖]

15

d_{ij} 的第二部分是分红或利息等收益[⊜]，然而有些资产没有这种收益，比如看涨、看跌期权和零息券等．

由于存在多种不同资产，并且每种资产可能有多种状态 d_{ij}. 我们可以用矩阵来表示这些数据．

对于所考虑的 N 种资产，它们的收益 d_{ij} 可以用矩阵 D 表示：

$$D=\begin{bmatrix} d_{11} & \cdots & d_{1K} \\ \vdots & \vdots & \vdots \\ d_{N1} & \cdots & d_{NK} \end{bmatrix} \tag{2.5}$$

有两种方式解释这个矩阵，矩阵 D 的每一行表示一个单位的给定证券在不同状态下的收益．而矩阵 D 的每一列表示给定状态下不同单位证券的收益．

如果每种资产的现价均不为 0，那么可以将 D 中第 i 行除以对应的 $s_i(t)$，得到第 i 个资产在不同状态下的总收益率．当收益与时间有关时，D 会带有一个表示时间 t 的下标．

2.2.4 证券投资组合

证券投资组合是一系列资产按照一定的比例组合而成的特定组合．用符号 θ_i 表示第 i 种资产所占的比例，确定了$\{\theta_i,\ i=1,\ \cdots,\ N\}$，就确定了一个证券投资组合．

若 θ_i 为正数，则表示做多(买入)该项资产，若 θ_i 为负数，则表示做空(卖出)该项资产. 若该证券投资组合中不包含某项资产，则其对应的 θ_i 为零．

⊖ 为实现资本的收益，交易者必须进行平仓.

⊜ 除了股票红利和债券息票，还有未来的投资．由于实行逐日盯市制度，合约持有人每天都会收到回报．然而，这些回报可能大于 0，也可能小于 0.

如果一个证券投资组合在所有状态下所获收益均相同，那么它的价值是完全可知的，即该证券投资组合是无风险的．

2.2.5 资产定价的一个简单例子

下面我们通过一个简单的模型来展示金融衍生品定价中的大多数重要结果．通过这个例子，首先，我们将解释衍生品定价背后的逻辑；其次，我们将引入在实际应用中实现这一逻辑所需要的数学工具．在本节中，模型的难度被刻意降低了，本章结尾将会讨论一般情况下的模型．

假设时间包括“现在”和“下一期”，并且它们的间隔时间长度为 Δ. 本书中 Δ 表示一段很短但并非无穷小的时间长度．

假定市场参与者只对三种资产感兴趣：

1）无风险资产，比如美国联邦短期国库券等．它到下一期开始时的总收益率为 $(1+r\Delta)$. [㊀] 由于它与状态无关，这种回报是“无风险”的．

2）标的资产，比如一只股票 $S(t)$ 等．假定在经历了很短的时间间隔 Δ 后，$S(t)$ 有两种可能取值，即有两种状态．由于两种状态下的收益不同，故 $S(t)$ 是有风险的．

3）衍生品，比如当前价格为 $C(t)$，执行价为 C_0，在“下一期”开始时到期的看涨期权．由于标的资产有两种可能取值，故该看涨期权也有两种可能取值．

这种设置非常简单——三种资产 $(N=3)$，两种可能状态 $(K=2)$. 第一种资产为无风险借贷，第二种资产为标的资产，第三种资产为期权．

这个例子有一定的现实意义．现实生活中（连续时间的）交易员可能考虑对某种特定期权进行持仓．如果所考虑的时间间隔非常短，相对于价格保持不变，这些资产的价格更可能发生上涨或下跌．因此，假设有两种可能状态是合理的．[㊁]

16

用前面提到的符号来表示上述信息，资产价格将由包含三个元素的向量 S_t 表示，

$$S_t = \begin{bmatrix} B(t) \\ S(t) \\ C(t) \end{bmatrix} \tag{2.6}$$

其中 $B(t)$ 表示无风险借贷，$S(t)$ 表示股票，$C(t)$ 表示以该股票为标的资产的看涨期权的价格. t 表示这些价格所对应的时间．

如前所述，收益将表示为矩阵 D_t. 三种资产意味着矩阵 D_t 有三行．同理，两种可能状态意味着矩阵 D_t 有两列．由于 $B(t)$ 是无风险借贷，故它在不同状态下的收益是相同的；$S(t)$ 是有风险性的，它的价格可能上涨为 $S_1(t+\Delta)$ 或下跌为 $S_2(t+\Delta)$；看涨期权的价格 $C(t)$ 将随着标的资产价格 $S(t)$ 的不同而发生改变．因此，D_t 可表示为：

㊀ 我们必须将无风险利率 r 和时间间隔 Δ 相乘，才能得到正确的回报率.

㊁ 事实上，我们后面会介绍，当 Δ 趋于 0 时，双状态过程可以很好地拟合维纳过程，或者说布朗运动.

$$D_t = \begin{bmatrix} (1+r\Delta)B(t) & (1+r\Delta)B(t) \\ S_1(t) & S_2(t) \\ C_1(t) & C_2(t) \end{bmatrix} \tag{2.7}$$

其中 r 为无风险的年收益率.

2.2.6 套利定理初探

本节将引入金融理论中用于计算衍生品公平市场价格的基本原则. 首先，我们需要进一步简化记号. 无风险借贷的数额由投资者决定，因此，我们假定

$$B(t) = 1 \tag{2.8}$$

前面，我们将时间间隔记为 Δ，在这里，我们令

$$\Delta = 1 \tag{2.9}$$

在这些特殊假定下，套利定理可以表述为：

定理 1 设 S_t，D_t 为式(2.6)及式(2.7)给定，并且假定两种状态都可能发生，

1）若存在正常数 ψ_1，ψ_2 使得资产价格满足

$$\begin{bmatrix} 1 \\ S(t) \\ C(t) \end{bmatrix} = \begin{bmatrix} (1+r\Delta) & (1+r\Delta) \\ S_1(t+\Delta) & S_2(t+\Delta) \\ C_1(t+\Delta) & C_2(t+\Delta) \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix} \tag{2.10}$$

则不存在套利机会；[㊀]

2）若不存在套利机会，则必存在满足(2.10)的正常数 ψ_1，ψ_2.

式(2.10)表示的关系称为无套利表达式，但这种关系在现实生活中是无法被观察到的. 事实上，$S_1(t+1)$ 及 $S_2(t+1)$ 就是标的资产的“可能”未来值，它们中只有被实现的状态对应的价格能被观察到.

那么 ψ_1，ψ_2 表示什么呢？根据套利定理表达式的第二行，如果一种资产在状态 1 下价格为 1，在状态 2 下价格为 0，那么

$$S(t) = (1)\ \psi_1 \tag{2.11}$$

因此，投资者能够在当期以 ψ_1 的投资取得下一期如果是状态 1 发生，则获得一单位收益、下一期如果是状态 2 发生，则没有收益的保险策略. 类似地，ψ_2 表示投资者愿意为取得在状态 2 下获得一单位收益、在状态 1 下没有收益的保险策略所愿意付出的代价. 显然，投资者在当期花费 $\psi_1+\psi_2$，就可以保证无论下一期何种状态发生，都可以获得 1 单位收益. 式(2.10)中的第一行同样可以实现这一策略. 根据它们的内涵，称 ψ_i ($i=1$，2)为状态价格.[㊁]

17

此处还有几点没有说清楚，事实上，人们自然会想到以下几个问题：

- 这个定理是如何推导出来的？
- 是什么使得 ψ_1，ψ_2 的存在性与无套利有关系？

㊀ 如果 $1+r>1$，则有 $\psi_1+\psi_2<1$，可以从矩阵方程中的第一行得到该关系式.

㊁ 一般来说，状态价格是与时间有关的；因此，必须带有下标 t. 为了记号上的简便，这里忽略了下标.

- 这个结果为什么与资产定价有关系？

我们暂时将前两个问题搁置起来，先回答第三个问题：什么类型的实际问题(如果有)可从ψ_1，ψ_2 的存在性得到呢？事实上，套利定理的表达式与资产定价有着非常重要的关系.

2.2.7 与套利定理相关的变量

套利定理为衍生品定价提供了一个非常完美且通用的方法.

再次考虑表达式：

$$\begin{bmatrix} 1 \\ S(t) \\ C(t) \end{bmatrix} = \begin{bmatrix} (1+r\Delta) & (1+r\Delta) \\ S_1(t+\Delta) & S_2(t+\Delta) \\ C_1(t+\Delta) & C_2(t+\Delta) \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix} \tag{2.12}$$

将第一行中的分红矩阵 D_t 与 ψ_1，ψ_2 组成的向量相乘，可以得到

$$1 = (1+r)\psi_1 + (1+r)\psi_2 \tag{2.13}$$

定义：

$$\begin{aligned} Q_1 &= (1+r)\psi_1 \\ Q_2 &= (1+r)\psi_2 \end{aligned} \tag{2.14}$$

由于状态价格均为正数，且由(2.13)知，

$$0 < Q_1 \leqslant 1$$

$$Q_2 + Q_1 = 1$$

因此，Q_i 均为正数，且和为 1. 如此我们可以将它们解释为所考虑的两种状态发生的概率. 之所以称为“解释”，是因为控制两种状态是否发生的真实概率通常与Q_1，Q_2 不同. Q_1、Q_2 由等式(2.14)所定义，并且与控制两种状态是否发生的真实概率无直接联系. 因此称$\{Q_1, Q_2\}$为风险调节综合概率.

2.2.8 综合概率的应用

当套利机会不存在时，风险调节概率存在. 换句话说，当“定价不合理的资产”不存在时才有正的常数$\{\psi_1, \psi_2\}$. 乘以无风险总收益 $1+r$ 就能确保$\{Q_1, Q_2\}$存在.⊖

风险调节概率在资产定价中的重要性源于：通过风险调节概率计算得到的资产期望值，在无风险利率 r 下的贴现值与资产现值相等.

再次考虑套利定理所推出的等式，可以发现表达式(2.10)包含了下面三个等式：

$$1 = (1+r)\psi_1 + (1+r)\psi_2 \tag{2.15}$$

$$S(t) = \psi_1 S_1(t+1) + \psi_2 S_2(t+1) \tag{2.16}$$

$$C(t) = \psi_1 C_1(t+1) + \psi_2 C_2(t+1) \tag{2.17}$$

将后两个等式右边乘以

18

⊖ 这是状态数量有限的情况. 在状态数量不可数时，我们需要添加额外条件，才能确保风险调整概率存在.

$$\frac{1+r}{1+r} \tag{2.18}$$

可得[1]

$$S(t)=\frac{1}{1+r}[(1+r)\psi_1 S_1(t+1)+(1+r)\psi_2 S_2(t+1)] \tag{2.19}$$

$$C(t)=\frac{1}{1+r}[(1+r)\psi_1 C_1(t+1)+(1+r)\psi_2 C_2(t+1)] \tag{2.20}$$

用对应的 $\widetilde{P}_i$ 代替$(1+r)\psi_i$($i=1$，2)可得

$$S(t)=\frac{1}{1+r}[Q_1 S_1(t+1)+Q_2 S_2(t+1)] \tag{2.21}$$

$$C(t)=\frac{1}{1+r}[Q_1 C_1(t+1)+Q_2 C_2(t+1)] \tag{2.22}$$

现在考虑如何解释上面两个式子．观察等式右端，括号外的 $1/(1+r)$是常利率 r 下的折现因子．而括号内的式子可以解释为资产的某种期望值．它是以“概率”Q_1、Q_2 作为权重，乘以各状态下 $S(t)$和 $C(t)$对应的价格再求和所得．因此，括号内的项可视为用风险调节概率计算得到的期望值．

尽管如此，等式(2.21)和(2.22)并不表示真正的期望值．但只要不存在套利，上面两式就成立，并且能够在实际计算中被运用．只要各状态的调节概率明确给出，我们就能在资产定价中使用这两个式子．

按上述方法解释 Q_1、Q_2 后，所考虑的任何资产，其现价等于未来资产贴现后的期望回报．此外，尽管资产是有风险的，我们使用了无风险概率进行贴现．

为了更好地理解风险调节概率的重要作用，我们来看一看用实际的概率计算时会发生什么．

首先，根据各状态真实发生的概率 P_1，P_2，我们可以得到实际的期望价格：

$$E^{\text{true}}[S(t+1)]=[P_1 S_1(t+1)+P_2 S_2(t+1)] \tag{2.23}$$

$$E^{\text{true}}[C(t+1)]=[P_1 C_1(t+1)+P_2 C_2(t+1)] \tag{2.24}$$

由于这些资产是有风险的，在用无风险利率折现时，通常[2]满足以下关系：

$$S(t)<\frac{1}{1+r}E^P[S(t+1)] \tag{2.25}$$

$$C(t)<\frac{1}{1+r}E^P[C(t+1)] \tag{2.26}$$

为了理解为什么会是不等式，我们考虑这样的情况：假如等式成立，即

$$S(t)=\frac{1}{1+r}E^P[S(t+1)] \tag{2.27}$$

---

[1] 只要 r 不等于-1，我们总可以这样做．

[2] 我们强调“通常”，因为风险资产的价格可能是与市场呈负相关的．这类资产风险溢价小于 0，被称为“负 beta”资产．

$$C(t)=\frac{1}{1+r}E^P[C(t+1)] \tag{2.28}$$

假定资产价格非零，上式可改写为

$$1+r=\frac{E^P[S(t+1)]}{S(t)} \tag{2.29}$$

$$1+r=\frac{E^P[C(t+1)]}{C(t)} \tag{2.30}$$

但这意味着风险资产的(真实)期望回报等于无风险资产的回报．这样就会产生矛盾，因为风险资产通常包含着对承受风险的奖励．如果没有对风险的补偿，也就没有投资者会持有风险资产．因此，对于风险资产，通常有

19

$$(1+r+\text{对 } S(t) \text{ 的风险溢价})=\frac{E^P[S(t+1)]}{S(t)} \tag{2.31}$$

$$(1+r+\text{对 } C(t) \text{ 的风险溢价})=\frac{E^P[C(t+1)]}{C(t)} \tag{2.32}$$

这意味着对于风险资产，通常有以下不等式[1]：

$$S(t)<\frac{1}{1+r}E^P[S(t+1)] \tag{2.33}$$

$$C(t)<\frac{1}{1+r}E^P[C(t+1)] \tag{2.34}$$

此时，我们就看到了无套利假设在资产定价中的重要性．无套利即意味着存在正整数 ψ_1、ψ_2，进而我们可以得到风险调整概率 Q_1、Q_2，并得出“综合”期望，满足

$$1+r=\frac{E^Q[S(t+1)]}{S(t)} \tag{2.35}$$

$$1+r=\frac{E^Q[C(t+1)]}{C(t)} \tag{2.36}$$

上述两式非常便于使用，它们将所有风险溢价内在化了．事实上，如果用综合期望计算，不需要另外计算任何风险溢价．因为相关贴现是通过无风险利率实现的．

2.2.9 鞅和下鞅

现在我们要引入几个金融资产定价的基本概念．在这里，我们先简单地给出部分名词的定义，更详细的内容将在后面的章节介绍．

假定在 t 时刻某人所拥有的信息记为 I_t，对任何 $s>0$，如果随机过程$\{X_t, t\geqslant 0\}$满足等式

$$E^P[X_{t+s}\mid I_t]=X_t \tag{2.37}$$

那么称随机过程$\{X_t, t\geqslant 0\}$为在概率 P 下的鞅．[2]

---

[1] 对“负 beta”资产而言，该不等式的不等号方向正好相反．

[2] 此外，还有一个鞅必须满足的条件，我们将在后面的章节中具体介绍．在此期间，我们假设这些条件期望是存在的，也就是说，它们的值有限．

如果对任何 $s>0$，有

$$E^{Q}[X_{t+s}|I_t] \geqslant X_t \tag{2.38}$$

那么称随机过程$\{X_t, t\geqslant 0\}$为在概率 Q 下的下鞅．

我们来解释一下为什么它们是基本概念．根据前几节的讨论可以发现，用无风险利率贴现的资产价格在实际概率下是下鞅，而在风险调节概率下是鞅．因此，只要我们采用风险调节概率，鞅理论中的工具都适用于我们的问题．由鞅等式有

$$X_t = E^{Q}[X_{t+s}|I_t] \tag{2.39}$$

其中 $s>0$，而 X_{t+s}定义如下：

$$X_{t+s} = \frac{S_{t+s}}{(1+r)^s} \tag{2.40}$$

这里 S_{t+s}和 r 分别表示证券的价格和无风险回报率．Q 是风险调节概率．根据以上分析，风险调节概率可以将所有(贴现)资产价格转化为鞅．

2.2.10 标准化

20

值得注意的是，在金融中，鞅通常和两个概念伴随出现．第一，鞅的定义通常要说明在哪个概率下．因此在 2.2.9 节，我们说贴现后的股票价格在风险调节概率 Q 下是鞅．第二，要注意 S_t 并不是鞅，将 S_t 通过使用$(1+r)^s$ 进行标准化后，才得到鞅．$(1+r)^s$ 表示 1 美元的无风险投资在 s 期后的复利回报，令

$$X_{t+s} = \frac{S_{t+s}}{(1+r)^s} \tag{2.41}$$

X_{t+s}才是鞅．

在本书的后半部分我们将研究一个有趣的问题．假如我们用 S_t 除以某些其他资产的值，如 C_t，将得到：

$$X^*_{t+s} = \frac{S_{t+s}}{C_{t+s}} \tag{2.42}$$

该比率是否会在某种概率 P^* 下成为鞅呢？答案是肯定的，相关的结论在对利率敏感的衍生工具定价中非常实用，更重要的是，它告诉我们，在选用某些资产进行标准化时，我们可以灵活地选取其他定义下的概率．这些内容将在第 17 章讨论．

2.2.11 回报率均衡

通过使用风险调节概率，我们可以得到另一个在资产定价中非常重要的结果．

在(2.10)给出的无套利表达式中，将等式两边除以资产现价，再各乘以$(1+r)$，并假设资产价格非零，我们得到

$$Q_1 \frac{S_1(t+1)}{S(t)} + Q_2 \frac{S_2(t+2)}{S(t)} = 1+r \tag{2.43}$$

$$Q_1 \frac{C_1(t+1)}{C(t)} + Q_2 \frac{C_2(t+2)}{C(t)} = 1+r \tag{2.44}$$

首先观察比率

$$\frac{S_1(t+1)}{S(t)}, \quad \frac{S_2(t+1)}{S(t)} \tag{2.45}$$

上述比率实际上分别是 $S(t)$ 在状态 1 和状态 2 下的总回报率．等式(2.43)和(2.44)说明如果用 Q_1、Q_2 来计算期望值，所有资产都将产生同样的期望回报．根据这一结论，在 Q_1 和 Q_2 下，所有期望回报率等于无风险利率 r[⊖]．这是一个在金融资产定价中被广泛利用的结论．

2.2.12 无套利条件

在前面的介绍中，我们已经得出无套利条件与 ψ_1、ψ_2 存在性的关系．将状态 1 和状态 2 的总回报率分别记为 $R_1(t+1)$ 和 $R_2(t+1)$：

$$R_1(t+1) = \frac{S_1(t+1)}{S(t)} \tag{2.46}$$

$$R_2(t+1) = \frac{S_2(t+1)}{S(t)} \tag{2.47}$$

现在我们可以用新符号改写(2.12)的前两行为：

$$1 = (1+r)\psi_1 + (1+r)\psi_2$$

$$1 = R_1\psi_1 + R_2\psi_2$$

将两式联立得到：

$$0 = ((1+r) - R_1)\psi_1 + ((1+r) - R_2)\psi_2 \tag{2.48}$$

我们希望 ψ_1，ψ_2 大于 0，要同时满足大于 0 的条件和上面的等式，当且仅当 $R_1<(1+r)<R_2$. 如果这个条件不成立，比如

$$(1+r) < R_1 < R_2$$

21

这意味着，只要以利率 r 借款，买入 $S(t)$，我们就可以确保获得正收益．因此，这就是一个套利机会．在此时，(2.48)的右边将为负，则等式不成立，故不存在大于 0 的 ψ_1、ψ_2. 同理，我们可以得到，如果

$$R_1 < R_2 < (1+r)$$

此时，我们只要卖空 $S(t)$ 并将所得收入以无风险利率 r 买入无风险债券，就能在将来获得正的收益．此时，等式(2.48)不再成立，因为右端项将恒为正．

因此可以看出，存在 $\psi_1>0$、$\psi_2>0$ 即意味着

$$R_1 < (1+r) < R_2$$

它是无套利的充要条件．

2.3 一个具体的例子

下面，我们讨论一个简单的例子．设股票的现值为

⊖ 在概率论中，“在 Q_1、Q_2 下”意味着“使用 Q_1、Q_2 作为概率测度”.

$$S_t = 100 \tag{2.49}$$

又设在下一时刻，股票只有两种可能取值：

$$S_1(t+1) = 100 \tag{2.50}$$

和

$$S_2(t+1) = 150 \tag{2.51}$$

所以，本题中下一时刻只有两种状态．

现有一种价格为 C，执行价格为 100 的看涨期权，期权在下一时刻到期．

最后，我们假定 1 单位的资金被用作无风险投资，回报率为 10%.

我们可以得到下列无套利表达式：

$$\begin{bmatrix} 1 \\ 100 \\ C \end{bmatrix} = \begin{bmatrix} 1.1 & 1.1 \\ 100 & 150 \\ 0 & 50 \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix} \tag{2.52}$$

注意到，看涨期权的价格 C 事先是不知道的．将 C 作为变量，我们将发现套利定理所起的作用．

2.3.1 问题 1：套利的可能性

将(2.52)的矩阵相乘，可得：

$$1 = 1.1\psi_1 + 1.1\psi_2 \tag{2.53}$$

$$100 = 100\psi_1 + 150\psi_2 \tag{2.54}$$

$$C = 0\psi_1 + 50\psi_2 \tag{2.55}$$

现假设从市场上观察到 $C=25$. 代入最后一个等式，可得：

$$50\psi_2 = 25 \tag{2.56}$$

也就是

$$\psi_2 = \frac{1}{2} \tag{2.57}$$

将其代入(2.54)，可得：

$$\psi_1 = 0.25 \tag{2.58}$$

但所得的 ψ_1、ψ_2 的值并不能使第一个式子成立：

$$1.1 \times 0.25 + 1.1 \times 1.5 \neq 1 \tag{2.59}$$

显然，对于已知的看涨期权价格 $C=25$，不可能找到正数 ψ_1，ψ_2 使得 3 个等式全部成立，故存在套利的可能性．

22

2.3.2 问题 2：无套利价格

仍考虑上一个问题：

$$\begin{bmatrix} 1 \\ 100 \\ C \end{bmatrix} = \begin{bmatrix} 1.1 & 1.1 \\ 100 & 150 \\ 0 & 50 \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix} \tag{2.60}$$

现在 C 的值未知，首先我们通过方程组前两个方程，可以求得 ψ_1，ψ_2 的唯一解：

$$\psi_1 = 0.7273, \quad \psi_2 = 0.1818 \tag{2.61}$$

再利用第三个方程，可求得

$$C = 9.09 \tag{2.62}$$

在这个价格下，套利不存在．

请注意，我们利用 ψ_1 和 ψ_2 得到了无套利价格 $C=9.09$. 如此看来，套利定理可以作为一种定价工具．

在这个例子中，ψ_i 恰好都是正数且解唯一，在其他问题中不一定都是这样．

2.3.3 一类不确定性

如果状态的数量大于 2 个，用该方法将无法求看涨期权的无套利价格．例如，考虑以下问题：

$$\begin{bmatrix} 1 \\ 100 \\ C \end{bmatrix} = \begin{bmatrix} 1.1 & 1.1 & 1.1 \\ 100 & 50 & 150 \\ 0 & 0 & 50 \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \end{bmatrix} \tag{2.63}$$

在该问题中，前两个方程不能得到 $\psi_i > 0$ 的唯一解(有无穷多解)，因而不能确定 C.

要求出看涨期权的无套利价格 C，必须挑选正确的 ψ_i. 原则上，它们可以结合潜在的经济均衡关系得出．

2.4 应用：二叉树模型

刚才讨论的例子非常简单，却为我们提供了一个著名的资产定价模型的方法．该模型就是格点模型[⊖]，二叉树模型是最简单的情形．

我们将简单展示该定价方法是如何利用套利定理的．

考虑价格为 C_t 的看涨期权和价格为 S_t 的标的资产．该期权执行价格为 C_0，在时刻 T 到期，$t<T$. 易知到期日期权的价值为

$$C_T = \max[S_T - C_0, 0] \tag{2.64}$$

首先我们将时间段$(T-t)$分割为 n 个更小的时间段，每一小段长为 Δ. 由于 Δ 足够小，在 Δ 时间内，S_t 的变化可简化为仅上升或仅下降．

所以我们假定在 Δ 内 S_t 可能上升 $\sigma\sqrt{\Delta}$ 或下降 $-\sigma\sqrt{\Delta}$：

$$S_{t+\Delta} = \begin{cases} S_t + \sigma\sqrt{\Delta} \\ S_t - \sigma\sqrt{\Delta} \end{cases} \tag{2.65}$$

显然，参数 σ 的大小决定了 $S_{t+\Delta}$ 在单位时间 Δ 内能变化多少．因此该参数被称作波动参数．假定 σ 已知，Δ 越小，S_t 变化也越少．

⊖ 也称为树形模型．

(2.65)描述的单步变化就是一个一期二叉树模型．图 2-1 展示了多期情况下股价变化的可能路径．

现假定已知 Δ 时间段内的无风险利率 r（常数），我们能否确定风险调整概率㊀呢？

由套利定理我们知道风险调整概率 Q_{up} 和 Q_{down} 必须满足

$$S_t = \frac{1}{1+r}\left[Q_{\mathrm{up}}(S_t + \sigma\sqrt{\Delta}) + Q_{\mathrm{down}}(S_t - \sigma\sqrt{\Delta})\right] \tag{2.66}$$

在该等式中，r，S_t，σ 和 Δ 是已知变量．前 3 个可以从市场观察到，Δ 由我们自己设定．因此，唯一未知的是 Q_{up}，而它也很容易求出来㊁．

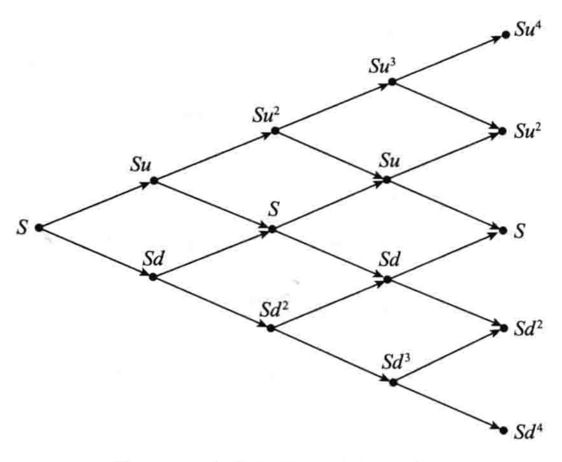


图 2-1 多期股价上下变动图

一旦完成上述步骤，则 Q_{up} 可以用来计算看涨期权的无套利价格．事实上，等式

$$C_t = \frac{1}{1+r}\left[Q_{\mathrm{up}} C_{t+\Delta}^{\mathrm{up}} + Q_{\mathrm{down}} C_{t+\Delta}^{\mathrm{down}}\right] \tag{2.67}$$

将 $t+\Delta$ 时刻的两个期权价值和 t 时刻的期权价值通过风险调整概率联系起来，Q_{up} 此时已知．要利用该等式，首先 $C_{t+\Delta}^{\mathrm{up}}$ 和 $C_{t+\Delta}^{\mathrm{down}}$ 必须已知．有了这些，我们才能计算 t 时刻看涨期权价值 C_t.

图 2-2 展示了多期情况下的期权价格 C_t 的二叉树图．除了到期日的期权值已知外，其余节点的 C_t 是未知的．事实上，通过 S_t 的二叉树图，我们可以确定 C_t 在到期日的价值，即利用边界条件

$$C_T = \max[S_T - C_0, 0] \tag{2.68}$$

一旦完成了这一步，我们就可以回过来使用

24

$$C_t = \frac{1}{1+r}\left[Q_{\mathrm{up}} C_{t+\Delta}^{\mathrm{up}} + Q_{\mathrm{down}} C_{t+\Delta}^{\mathrm{down}}\right] \tag{2.69}$$

算出倒数第二期所有节点的价格．重复该过程，最终就能达到起始点，即得期权的当前价值．

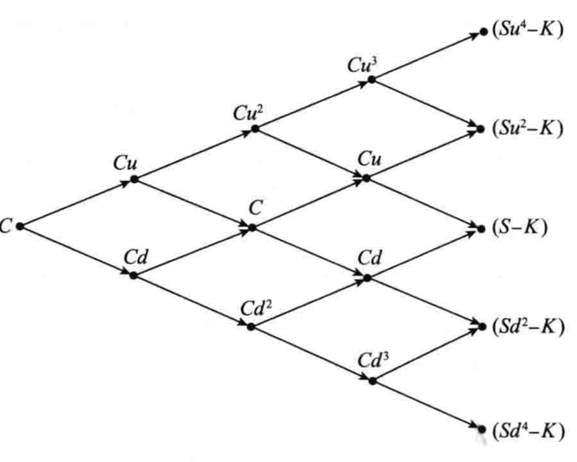


图 2-2 期权价格的多期二叉树图

㊀ 在本书后半部分，我们将放宽 r 为常数这一假设．现在我们暂时保留该假设．

㊁ 记住 $Q_{\mathrm{down}} = 1 - Q_{\mathrm{up}}$.

因此，该过程先根据 S_t 的变化规律向前进，直到到期日，再确定到期日的期权价格．然后利用风险调整概率和边界条件，沿树形倒退，最终得到当前期权价格 C_t.

正是有了套利定理和鞅作为工具，我们才能够计算风险调整概率 Q_{up} 和 Q_{down}.

在上述过程中，图 2-1 近似地给出了在 $T-t$ 时段内 S_t 的所有可能路径．图 2-2 则给出了基于 S_t 的看涨期权价格的所有可能路径．Δ 越小，上述二叉树就越接近真实的移动路径.

2.5 红利与外币

本节我们将对本章介绍的双状态模型进行调整，来适应两种在实际中更常见也更复杂的情况．第一种是带中期支付的股票期权，例如股利和息票．许多证券都可能在期权的到期日前作此类支付，这些支付会对定价公式产生影响．第二种是标的资产的价格由外币决定的期权，这时定价公式也会略有微调．

2.5.1 有分红的情况

在 2.3 节所讨论的问题中，我们添加支付红利的环节，即支付 $S_{t+\Delta}$ 的 d_t%的股利．注意两点，第一，股利并不是一次付清的，而是在每个 $t+\Delta$ 时刻支付一定比例；第二，股利支付率的下标是 t 而不是 $t+\Delta$，故在 t 时刻我们就知道 d_t 的大小，所以给定信息集 I_t 后，d_t 不再是一个随机变量．

25

(2.10)中的简单模型现在演化成了如下形式：

$$\begin{bmatrix} 1 \\ S_t \\ C_t \end{bmatrix} = \begin{bmatrix} B^u_{t+\Delta} & B^d_{t+\Delta} \\ S^u_{t+\Delta} + d_t S^u_{t+\Delta} & S^d_{t+\Delta} + d_t S^d_{t+\Delta} \\ C^u_{t+\Delta} & C^d_{t+\Delta} \end{bmatrix} \times \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix}$$

其中 B、S、C 含义不变，分别表示储蓄额、股价和看涨期权价格，为了适应 2.5 节的新环境，这些记号作了一些改变．

我们能否把 2.3 节的步骤用在这里呢？答案是肯定的，只需略微改动，我们可以沿用上述思路得到以下等式：

$$S = \frac{1+d}{1+r}[S^u Q_{\text{up}} + S^d Q_{\text{down}}] \tag{2.70}$$

$$C = \frac{1}{1+r}[C^u Q_{\text{up}} + C^d Q_{\text{down}}] \tag{2.71}$$

其中 Q 是风险中性概率，且式中没有时间下标．可以看到，第一个等式发生了改变，而第二个等式没有变化．分析可知，每个 Δ 之后股票持有者可获得 d%的股利，对支付红利的资产进行无风险折现，必须乘以$(1+d)/(1+r)$而非 $1/(1+r)$. 另外值得强调的是衍生品的折现方式并不改变．

现在进行如下变形：

$$\frac{1+r}{1+d} = \left[\frac{S^u Q_{\text{up}} + S^d Q_{\text{down}}}{S}\right]$$

这意味着无风险条件下的期望回报率是

$$E^{Q}\left[\frac{S_{t+\Delta}}{S}\right]=\frac{1+r\Delta}{1+d\Delta}$$

如果 d、r 为年利率，且都较小时，那么有下列一阶近似：

$$\frac{1+r\Delta}{1+q\Delta}\approx 1+(r-d)\Delta$$

将其代入之前的等式，得

$$E^{Q}\left[\frac{S_{t+\Delta}}{S_t}\right]\approx 1+(r-d)\Delta$$

或

$$E^{Q}[S_{t+\Delta}]\approx S_t+(r-d)S_t\Delta$$

增加一个随机、不可预测的成分 $\sigma S_t\Delta W_{t+\Delta}$，则得[⊖]：

$$S_{t+\Delta}\approx S_t+(r-d)S_t\Delta+\sigma S_t\Delta W_{t+\Delta}$$

我们可以对上式作如下解释：如果令 Δ 趋于 0，时间连续，那么 $\mathrm{d}S_t$ 的漂移项——标的资产价格变化期望值的项，可以用 $(r-t)S_t\mathrm{d}t$ 表示，从而得到：

$$\mathrm{d}S_t=(r-d)S_t\mathrm{d}t+\sigma S_t\mathrm{d}W_t$$

其中 $\mathrm{d}t$ 表示一段无穷小的时间.

接下来我们用相同方法处理 C_t 的等式(2.71)：

$$C=\frac{1}{1+r}[C^{u}Q_{\mathrm{up}}+C^{d}Q_{\mathrm{down}}]$$

可得：

26

$$E^{Q}\left[\frac{C_{t+\Delta}}{C_t}\right]1+r\Delta$$

由此，我们可以看到即使标的股票支付红利，其看涨期权的风险中性期望回报和无风险折现与无红利时相同．因此，在风险中性世界中，无论是否支付股利，都使用同一个折现因子，即 $1/(1+r)$.

换句话说：

- S_t 和 C_t 在 Δ 时间段无风险概率 Q 下的期望回报率分别为：

$$E^{Q}\left[\frac{S_{t+\Delta}}{S_t}\right]=\frac{1+r\Delta}{1+d\Delta}\approx 1+(r-d)\Delta$$

$$E^{Q}\left[\frac{C_{t+\Delta}}{C_t}\right]=1+r\Delta\approx 1+r\Delta$$

这些在公式很小的改动，在实际的定价计算中会产生巨大的区别．下面关于外币的问题也有类似的结果.

⊖ 我们将在后面的章节具体学习此类随机微分方程.

2.5.2 外币的情况

现在我们修改初始模型，假定可以通过外币储蓄进行投资．

具体而言，假定我们能用 e_t 单位本币购买 1 单位外币(美元为本币)，也就是说，t 时刻汇率为 e_t，外币利率已知，为 r^f.

投资机会和 Δ 后的收益可以用以下方程组概括：

$$\begin{bmatrix} 1 \\ 1 \\ C_t \end{bmatrix} = \begin{bmatrix} 1+r & 1+r \\ \frac{e_{t+\Delta}^u}{e_t}(1+r^f) & \frac{e_{t+\Delta}^d}{e_t}(1+r^f) \\ C_{t+\Delta}^u & C_{t+\Delta}^d \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix}$$

其中 C_t 是建立在汇率为 e_t 的基础上的看涨期权价格，执行价格为 K[㊀].

用与有红利时的方法一样，可以得到下列等式[㊁]：

$$e = \frac{1}{1+r}[e^u Q_{\text{up}} + e^d Q_{\text{down}}]$$

$$C = \frac{1}{1+r}[C^u Q_{\text{up}} + C^d Q_{\text{down}}]$$

同样，可以发现第一个等式发生了变化而第二个不变．所以，当我们处理与外币相关的资产时，若其 Δ 内的回报率为 r^f，该资产的无风险折现系数是 $(1+r)/(1+r^f)$.

如果 r^f 较小，我们可以作以下近似：

$$\frac{(1+r\Delta)}{(1+r^f\Delta)} \approx 1 + (r - r^f)\Delta$$

我们再一次能得到不同的结果：

- 在概率 Q 下，e_t 和 C 的期望回报率是不同的：

$$E^Q\left[\frac{e_{t+\Delta}}{e_t}\right] \approx 1 + (r - r^f)\Delta$$

$$E^Q\left[\frac{C_{t+\Delta}}{C_t}\right] \approx 1 + r\Delta$$

由上一节可知，如果令 Δ 趋于 0，dC_t 的漂移项可用 $rC_t dt$ 表示．但外汇资产 de_t 的漂移项只能是 $(r-r^f)e_t dt$.

27

2.6 推广

到目前为止，我们讨论的模型都非常简单，通常还不足以对现实世界的金融资产进行定价．下面我们来对前面的内容作一些推广．

㊀ 这里 K 表示汇率 e_t 的执行价格．如果汇率在 $t+\Delta$ 时刻超过 K，该看涨期权的购买方将收到 $(e_{t+\Delta}-K)N$ 的差价.

㊁ 通常，为方便起见，我们忽略时间下标．

2.6.1 时间指标

在此之前，我们考虑的时间都是离散的（$t=1$，2，3…），在时间连续的定价模型中，这一点将发生改变．我们假设 t 是连续的：

$$t \in [0,\infty) \tag{2.72}$$

在此之前，我们用 Δ 来代表一个很小的时间段，而现在我们能够用 $\mathrm{d}t$ 代表一个无穷小的时间段．

2.6.2 状态

在连续时间下，资产在未来可能的取值不局限于两种，可能会有无穷多种可能性，甚至构成连续的区间．

为了描述这种思想，我们需要引入随机微分方程．例如，上一节的例子中，我们用以下式子来表示 S_t 的增量 $\mathrm{d}S_t$：

$$\mathrm{d}S_t = \mu_t S_t \mathrm{d}t + \sigma_t S_t \mu \mathrm{d}W_t \tag{2.73}$$

其中 $\mathrm{d}S_t$ 表示证券价格在一个无穷小时段上的变化值，$\mu_t S_t \mathrm{d}t$ 是在无穷小的时间段 $\mathrm{d}t$ 中可预测的变化量，而 $\sigma_t S_t \mathrm{d}W_t$ 是无穷小时段上不可预测的随机波动．

由于随机微分方程是一个复杂的课题，其中的概念需要逐步引入．

2.6.3 折现

由于模型中时间连续化，价值的折现方式也要发生改变．事实上，如果将 t 视为连续时间，那么 Δ 时间段的折现因子就是

$$\mathrm{e}^{-r\Delta} \tag{2.74}$$

r 是按连续复利计算的利率．假如存在股利或外币，则 r 也需按照 2.5 节中所述方法进行修正．

2.7 小结：资产定价方法

套利定理为市场上金融产品定价提供了一种有效的方法．现对其主要步骤概括如下：

1）利用模型估计其标的资产可能的价格变化路径．

2）通过标的资产的价格，计算到期日或其他边界上的衍生品价格．

3）得到风险调整概率．

4）从到期日向前倒推，利用风险调整概率计算期望回报．

5）用无风险回报率对期望收益折现．

要使用这种定价方法，必须熟悉以下几个数学工具．

第一，要谨慎定义关于时间的概念．如果要处理无穷小时间段内资产价格的变化，就需要进行连续时间分析．

第二，我们要处理无穷小时间段内的随机量．无穷小的时间段内，概率、期望、平均

28

值、波动率等概念也要谨慎定义，这需要随机微积分方面的知识．我们将讨论得到微积分中主要结果的假设背后的实际背景．

第三，我们需要理解如何获得风险调整概率和正确的折现因子．Girsanov 定理介绍了使用风险调整概率所需要的条件．该定理还给出了这些概率分布的形式．

第四，无论在 Girsanov 定理中，还是在理解风险中性世界的过程中，鞅的概念都很重要．

第五，还有一个问题，如何度量随机变量在两个不同时间之间变化的大小．在标准计算中，需要运用微分方程解决．而在随机环境下，就要运用随机微分方程．

毫无疑问，要解决这些问题，你必须理解基本概念和掌握计算方面的知识．主要有三方面：(1)衍生品的概念；(2)积分的知识；(3)泰勒展开式的知识．

2.8 参考阅读

本章对套利定理作了简单介绍．Ingersoll(1987)提供了更详细且易懂的介绍，即使是初学者也很容易理解．数学能力较强的读者可以参考 Duffie(1996)的著作．Harrison 和 Kreps(1979)的文章也可以参阅．其他相关资料可以在 Harrison 和 Pliska(1981)的著作中找到．Musiela 和 Rutkowski(1997)的第一章写得极好，且非常适合学完本章后阅读．

2.9 附录：套利定理的一般形式

根据套利定理，如果不存在套利可能性，就存在状态价格$\{\psi_i\}$，使得每种资产的当前价格可以写成其未来价格的线性组合．反之亦然，如果存在满足条件的状态价格，就不存在套利机会．

本节将介绍套利定理的一般形式．我们定义以下符号：

- 定义回报矩阵 D：

$$D_t = \begin{bmatrix} d_{11} & \cdots & d_{1K} \\ \vdots & \vdots & \vdots \\ d_{N1} & \cdots & d_{NK} \end{bmatrix} \tag{2.75}$$

  N 等于证券种类的数目，K 等于未来状态的数目．

- 定义投资组合 θ，向量中的 N 个值与 N 种资产一一对应．

$$\theta = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_N \end{bmatrix} \tag{2.76}$$

  在操作者眼里，θ 对应某时刻各资产的配比．将 θ 与 S_t 相乘，即可得到投资组合的价格：

$$S'_t\theta = \sum_{i=1}^{N} S_i(t)\theta_i \tag{2.77}$$

  这就是 t 时刻投资组合 θ 的总价值．

• j 状态下投资组合 θ 的总收益是 $\sum_{i=1}^{N} d_{ij}\theta_i$，用矩阵表示为：

$$D'\theta = \begin{bmatrix} d_{11} & \cdots & d_{N1} \\ \vdots & \vdots & \vdots \\ d_{1K} & \cdots & d_{NK} \end{bmatrix} \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_N \end{bmatrix} \tag{2.78}$$

29

• 现在可以定义一个套利组合：

定义 5 只要满足以下任意一个条件，θ 即为套利组合：

1) $S'\theta \leqslant 0$ 且 $D'\theta > 0$

2) $S'\theta < 0$ 且 $D'\theta \geqslant 0$

若满足上述条件 1，投资组合 θ 即能在任何情况下都获得正收益，且无需购买其他资产. 若满足条件 2，该组合能在当前时刻作负的支出，未来获得非负回报.

以下定理是之前讨论的套利定理的一般形式.

定理 2 假如不存在套利机会，则存在 $\psi > 0$ 使得

$$S = D\psi \tag{2.79}$$

假如存在 $\psi > 0$ 使得(2.79)成立，则没有套利机会.

这意味着无套利世界中存在 ψ_i 满足

$$\begin{bmatrix} S_1 \\ \vdots \\ S_N \end{bmatrix} = \begin{bmatrix} d_{11} & \cdots & d_{1K} \\ \vdots & \vdots & \vdots \\ d_{N1} & \cdots & d_{NK} \end{bmatrix} \begin{bmatrix} \psi_1 \\ \vdots \\ \psi_K \end{bmatrix} \tag{2.80}$$

由于对任何 i，都有 $\psi_i > 0$，即每一种未来状态发生的概率都为正.

先考虑一个特殊的回报矩阵：

$$D = \begin{bmatrix} 1 & \cdots & 1 \\ d_{21} & \cdots & d_{2K} \\ \vdots & \vdots & \vdots \\ d_{N1} & \cdots & d_{NK} \end{bmatrix} \tag{2.81}$$

在矩阵 D 中，第一行是常数且为 1，这意味着第一种资产在将来只有 1 种状态. 也就是说，第一种资产是无风险的.

利用套利定理，将 D 的第一行与状态价格向量相乘，得

$$S = \psi_1 + \cdots + \psi_K \tag{2.82}$$

定义

$$\sum_{i=1}^{K} \psi_i = \psi_0$$

ψ_0 可视为无风险借款中 1 单位还款所对应的本金.

2.10 习题

1. 假如你有一份无红利的股票 S_t 和对应的看涨期权 C_t，未来只有两种可能的状态：

$$S_t = \begin{cases} 320 & 若\ u\ 发生 \\ 290 & 若\ d\ 发生 \end{cases}$$

两种状态的真实概率是$\{P_{up}=0.5, P_{down}=0.5\}$. 股票现价 $S_t=280$，年利率为常数，$r=5\%$. 时间是离散的，$\Delta=3$ 个月. 执行价格 $K=280$，到期日 $t+\Delta$.

(a) 找到风险中性概率 Q，使得在 Q 下资产价格构成鞅，折现率取无风险利率.

(b) 利用上题结论以及 $C_t=\frac{1}{1+r\Delta}E^Q[C_{t+\Delta}]$，计算期权价格.

(c) 用 S_t 进行标准化，并找到新测度 $\widetilde{P}$，使得标准化后的变量是鞅.

(d) 写出用 S_t 标准化后的鞅等式.

(e) 用 $\widetilde{P}$ 计算期权的公平市场价格.

(f) 能否说公平市场价格和鞅的选择无关?

(g) 如何通过两种概率得到同样的无套利价格?

(h) 最后，期权的风险溢价是多少? 实际中我们能否计算? 为什么?

2. 某经济系统有 2 种未来状态、4 种资产. 现给出其中 3 种证券在各状态下的价格及股利:

	价格		股利	
	状态 1	状态 2	状态 1	状态 2
证券 A	120	70	4	1
证券 B	80	60	3	1
证券 C	90	150	2	10

A、B、C 的现价分别为 100、70、180.

(a) 这三种证券的现价是否是无套利的?

(b) 如果不是，那我们可以构造怎样的套利组合?

(c) 求这三种证券的无套利价格.

(d) 现假定我们引入第四种证券，是下一时间段关于 B 的期货合约. 它的价格应该是多少?

(e) 假定现有执行价格 $K=125$ 的关于 C 的看跌期权. 该期权在第二个时间段到期. 它的无套利价格是多少?

3. 考虑股票 S_t 和对应的普通平价看跌期权. 期权在 $t+\Delta$ 到期，其中 Δ 代表一段较小的时间，在 t 时刻，S_t 有两种可能的走向，可能上升到 $S^u_{t+\Delta}$，或下降到 $S^d_{t+\Delta}$. 交易者可以以年利率 r 无风险地借入或借出资金.

(a) 利用套利定理，写出以上信息所对应的向量方程组，其中包含 3 个等式、2 种状态以及 S_t、C_t 的无套利价格.

(b) 画出 S_t 的两期二叉树. 假定每一个节点都是无套利的，它包含了多少个(a)中的三等式系统?

(c) 能否写出一个包含 3 个等式、4 种状态的系统，使之与(b)中的二叉树对应?

(d) 如何判断各状态的状态价格是没有矛盾的?

4. 建立股票 S_t 的四期二叉树，其上升、下降参数如下：$u=1.15$，$d=1/u$. 每期时间间隔为 1 个月，即 $\Delta=1$，则 S_t 的动态变化如下:

$$S^u_{t+\Delta}=S_t u,\ S^d_{t+\Delta}=S_t d$$

上升和下降代表每个节点的两种未来状态. 假定单位时间为月，相应的欧式看涨期权的到期时间为 $t=4$. 该股票无股利，且市场参与者期望该股票每年上涨 15%，无风险利率 r 恒为 5%.

(a) 根据以上信息，假定 S_t 服从对数正态分布，年波动率约为多少?

(b) 计算并写出 S_t、C_t 的四期二叉树.

31

(c) 计算 $t=0$ 时刻期权的无套利价格 C_0.

5. 关于股票 S_t 已知以下信息:

- 现价为 102.
- 年波动率为 30%.
- 现货利率为 5%, 未来 3 个月保持不变.

现将单位时间定为 1 个月, 利用二叉树来研究 S_t 的动态变化.

(a) 考虑一个关于 S_t 的欧式看涨期权. 期权执行价格为 $K=120$, 期限为 3 个月. 利用 S_t 和无风险借贷 B_t, 构造一个能够替代该期权的投资组合.

(b) 利用替代原期权的投资组合, 为看涨期权定价.

(c) 假如你在场外卖出 100 份这种看涨期权, 你将如何防止损失? 请具体作答.

(d) 假如该看涨期权的市场价为 5, 请建立一个有套利的投资组合.

6. 假设已知以下信息:

- 无风险年利率为 $r=6\%$.
- 股票价格满足 $S_t-S_{t-1}=\mu S_t+\sigma S_t\varepsilon_t$, 其中 ε_t 是独立的二项过程, 取值如下:

$$\varepsilon_t=\begin{cases}+1 & \text{概率为 } p\\ -1 & \text{概率为 } 1-p\end{cases}$$

  其中 $0<p<1$ 为参数.
- 年波动率为 12%.
- 该股票无股利, 股票现价为 100.

现考虑以下问题:

(a) 假如 μ 等于无风险利率:

$$\mu=r$$

且 S_t 是无套利价格. 那么 p 等于多少?

(b) 假如 $p=1/3$, 那么无套利价格 S_t 是否存在?

(c) 现假定 $\mu=r+$ 风险溢价, 在这些条件下, p 和 ε_t 代表什么?

(d) 能够确定 p 的值吗?

7. 使用上一题中的数据. 现在需要你计算基于股票 S_t 的欧式看涨期权价格的当前值. 该期权执行价格为 100, 期限为 200 天.

(a) 确定时间间隔 Δ, 使二叉树为 5 期.

(b) u 和 d 的值分别是多少?

(c) 每一期上升的概率是多少?

(d) 求股票价格 S_t 的二叉树.

(e) 求看涨期权 C_t 的二叉树.

8. 假定年利率 r 等于 1%. 某股票的年波动率为 10%. 利用二叉树为一份 3 年期的欧式看涨期权定价. 若该股票价格超过 55, 该期权会支付 1 美元. 股票现价为 50 美元. 画出 3 期二叉树, 每一期的时间间隔为 1 年.

32

9. 用 Matlab 编写程序, 解决上题中的问题.

第3章 确定性微积分回顾

3.1 引言

衍生品资产的数学研究通常假定时间是连续变化的，这就使得新信息是连续出现的，决策者必须面对各种新事件带来的无穷小时间内的变化．所以我们需要随机微积分，来处理衍生品定价中随机变量在无穷小时间段内的变化．

随机微积分是由许多运算规则构成的一个完整体系，与常规的微积分有许多区别．

刚接触时，学习者可能觉得随机微积分过于抽象，然而这种第一印象并不正确．连续时间下的金融产品研究较离散时间的情况有所加深也有所简化．市场参与者经过一段时间的学习和实践后，往往会觉得用连续时间的工具比离散时间更为得心应手．

事实上，连续时间下可以得到一些离散时间下无法实现的结果．随机微积分为金融研究打开了一扇大门．例如，连续时间允许投资组合中各成分的比例为无穷小时间内的变化，这样，用“简单”投资组合替代非线性资产成为可能．要替代一份期权，可能会用到标的资产和无风险借款，在离散时间下是无法做到精确替代的．㊀

3.1.1 信息流

与常规微积分相比，随机微积分更适合研究金融市场上的信息流．

例如，“时间间隔”的选取在不同交易日可能是不同的．市场的波动程度时强时弱．若波动率经常变化，观察期也需要加以调整，上一章中的 Δ 即为单位观察期．

另外，证券定价的数量方法往往计算量很大，故研究者需要视波动率大小选择时间间隔，在精确度与计算效率之间寻求平衡．但这种近似方法在连续时间下就不用考虑，我们将用随机微积分来处理这些模型．

3.1.2 对随机行为建模

随机微积分有一个技术上的优点，即对于复杂的随机变量，一旦在连续时间下，基于无穷小时间段研究，其结构将非常简单．例如，将时间间隔记作 dt，如果 dt 无穷小，我们就有理由认为资产价格在下一个瞬间只可能有两种变化，即或升或降．

在一定条件下，这种无穷小时间段 dt 下的二元结构是一种很好的近似方法，但这在时间间隔为 Δ 的离散时间下没有必要．㊁

此外，随机微积分的主要工具——伊藤积分在金融市场中比常规微积分中的黎曼积分更

㊀ 除非未来的状态是离散的，也就是说，标的资产在未来只有若干种可能的值．

㊁ 二点分布随机变量只有两个可能的取值，相比于有无穷多个取值的随机变量，它使工作大大简化了．

加适用.

这些都是建立一种新型微积分的理由．但在此之前，复习一下常规微积分是非常必要的. 尽管随机微积分的规则不同，但二者内在逻辑是相同的．

- 有时我们要计算一个变量的随机变化会对另一个变量产生什么影响，所以要用到各种函数的求导．
- 有时我们要计算随机增量的和，所以需要(随机)积分．
- 有时我们要用更简单的函数近似一个套利函数，这就要用到(随机)泰勒公式．
- 最后，我们要对连续时间下随机变量的行为动态建模，这就需要随机微分方程．

3.2 一些常规微积分工具

本节我们将回顾常规确定性微积分的主要概念，即使读者熟悉这些基础概念，也能够从本节的例子中有所收获，这些例子也反映了常规微积分在处理随机变量时的局限性．

34

3.3 函数

假设 A，B 是两个集合，f 是一种规则，使得任意一个 A 中的元素，都能通过 f 对应到 B 中的唯一一个元素，[⊖]这样的规则 f 叫做函数或者映射．在数学分析中，函数记作

$$f:A \to B \tag{3.1}$$

或

$$y = f(x)\ ,x \in A \tag{3.2}$$

假如 B 是实数集，我们说 f 是实值函数，记作

$$f:A \to \mathbf{R} \tag{3.3}$$

假如 A 和 B 都是函数集，那么 f 的功能是将一个函数对应到另一个函数，f 称为算子．

大多数读者应该都熟悉常规的函数记号，可是很少有读者接触过随机函数．

3.3.1 随机函数

在函数

$$y = f(x)\ ,\quad x \in A \tag{3.4}$$

中，只要知道 x 的值，就能得到 y 的值．通常假定 y 是实数，现在作以下重要调整．

存在集合 Ω，其中元素 ω 代表状态．函数 f 依赖于 $x\in\mathbf{R}$ 和 $\omega\in\Omega$：

$$f:\mathbf{R} \times \Omega \to \mathbf{R} \tag{3.5}$$

或

$$y = f(x,\omega),\quad x \in \mathbf{R},\quad \omega \in \Omega \tag{3.6}$$

其中记号 $\mathbf{R}\times\Omega$ 表示 $f(\cdot)$中有 2 个变量，一个来自集合 Ω，另一个来自集合 $\mathbf{R}$.

函数 $f(x,\omega)$具有以下性质：给定 $\omega\in\Omega$，$f(\cdot,\omega)$将成为只依赖 x 的函数．因此给定不

⊖ 集合 A 叫做定义域，集合 B 叫做 f 的值域．

同的 $\omega\in\Omega$，我们能得到不同的关于 x 的函数．图 3-1 即给定 2 个不同 ω 时函数关于 x 的变化情况．

若 x 代表时间，我们可以将 $f(x,\omega_1)$、$f(x,\omega_2)$ 理解为不同状态下的两种变化轨迹．

若 ω 代表潜在的随机因素，函数 $f(x,\omega)$ 就可以被称作随机函数，随机函数也叫随机过程．在随机过程中，x 代表时间，通常令 $x\geqslant 0$.

请注意，随机过程中的随机因素是直接决定整条路径的，而不是只决定某个时刻的某个值．也就是说，有许多轨迹构成一个路径集，所谓随机就是从中挑选一条．选择了一种状态 ω，就确定了整条路径．

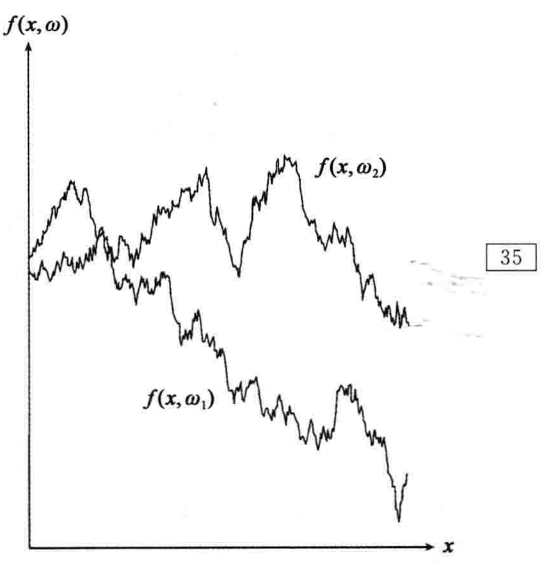


图 3-1 函数 $f(x,\omega)$ 的两个不同 ω 值对应的图像

35

3.3.2 函数举例

我们将用到许多重要的函数，下面简单介绍一下.

1. 指数函数

无穷项的和

$$1+1+\frac{1}{2!}+\frac{1}{3!}+\cdots+\frac{1}{n!}+\cdots \tag{3.7}$$

当 $n\to\infty$ 时，收敛于一个介于 2 和 3 之间的无理数．我们将这个数记作 e. 将 e 作为底数、x 作为指数，即可得到指数函数

$$y=\mathrm{e}^x,\quad x\in\mathbf{R} \tag{3.8}$$

该函数通常用于在连续时间内对资产价格进行折现．

指数函数有一个重要的性质：该函数是无穷次可微的．令 $y=\mathrm{e}^{f(x)}$，以下运算可进行无穷次递归运行：

$$\frac{\mathrm{d}y}{\mathrm{d}x}=\mathrm{e}^{f(x)}\frac{\mathrm{d}f(x)}{\mathrm{d}x} \tag{3.9}$$

指数函数还有一个有趣的乘法运算性质：

$$\mathrm{e}^x\mathrm{e}^z=\mathrm{e}^{x+z} \tag{3.10}$$

最后，如果 x 是随机变量，那么 $y=\mathrm{e}^x$ 也是随机的．

2. 对数函数

对数函数是指数函数的反函数．已知

$$y=\mathrm{e}^x,\quad x\in\mathbf{R} \tag{3.11}$$

y 的自然对数表达为

$$\ln(y)=x,\quad y>0 \tag{3.12}$$

有时候从业者的研究对象是资产价格的对数，请注意，这里 y 恒大于 0，x 则无限制．因此资产价格的对数取值范围是负无穷到正无穷．

3. 有界变差函数

以下内容将会在后面的章节多次用到．

假定将一段时间间隔定义为$[0,T]$，将其分为 n 个小区间并将端点记作 $t_i(i=1, \cdots, n)$：

$$0=t_0<t_1<\cdots<t_n=T \tag{3.13}$$

(t_i-t_{i-1})即为第 i 个小区间的长度．

现考虑定义在$[0, T]$上的函数 $f(t)$：

$$f:[0,T]\rightarrow \mathbf{R} \tag{3.14}$$

建立级数

$$\sum_{i=1}^{n}|f(t_i)-f(t_{i-1})| \tag{3.15}$$

显然，对区间$[0,T]$用不同方式切割，都能得到这样的求和．分割的方式有无数种，求和的结果也有无数种，若这些和所有的值有上界，则称函数 $f(\cdot)$是有界变差函数．这意味着

36

$$V_0=\max\sum_{i=1}^{n}|f(t_i)-f(t_{i-1})|<\infty \tag{3.16}$$

其中最大值是对所有可能的分法而言的．因此，V_0 是所有可能的变差的上界，且是有限的，称 V_0 是 f 在$[0,T]$上的全变差．粗略地讲，V_0 衡量了 $f(\cdot)$的轨迹从 0 到 T 的总长度．

因此，有界变差函数并非无规律可循，事实上，任何“光滑的”函数都是有界变差函数．㊀

4. 一个例子

考虑函数

$$f(t)=\begin{cases} t\sin\left(\dfrac{\pi}{t}\right) & 当\ 0<t\leqslant 1\ 时 \\ 0 & 当\ t=0\ 时 \end{cases} \tag{3.17}$$

可以证明该函数不是有界变差函数㊁．

图 3-2 是 $f(t)$的函数图像，可以看到当$t\rightarrow 0$时，f 变得非常“不规则”．

㊀ 可证明如果函数在$[0,T]$上处处可导，那么该函数是有界变差函数．

㊁ 为严格证明，选取分割

$$0<\frac{2}{2n+1}<\frac{2}{2n-1}<\cdots<\frac{2}{5}<\frac{2}{3}<1 \tag{3.18}$$

该分割上的变差为

$$\sum_{i=1}^{n}|f(t_i)-f(t_{i-1})|=4\frac{1}{3}+\frac{1}{5}+\frac{1}{7}+\cdots+\frac{1}{2n+1} \tag{3.19}$$

等式右端当 $n\rightarrow\infty$时趋于无穷大．

有界变差的概念在我们后面的学习中非常重要．原因之一是：资产价格在连续时间下有些部分不可预测，无论我们如何精细地分割时间区间，还是会有些部分不可预测，但这意味着资产价格的轨迹一定是非常不规则的．

后面我们将会看到，某些资产价格轨迹在连续时间下是非有界变差的．

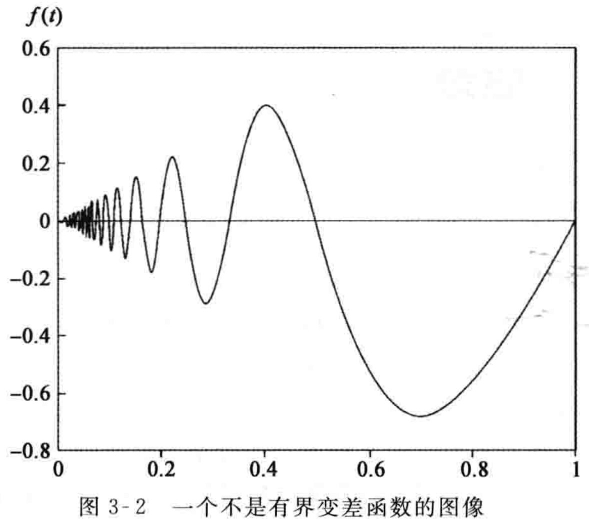


图 3-2　一个不是有界变差函数的图像

3.4　收敛和极限

假设已知数列

$$x_0, x_1, \cdots, x_n, \cdots \tag{3.20}$$

其中 x_n 是一个随着 n 增加而变化的对象，这些“对象”可以是一列数、一列函数，或一列算子，还有关键的一点是我们是依次观察 x_n 的．

收敛所关注的是当 $n \to \infty$ 时，x_n 最终会趋于什么值．假定 x_n 为实数，我们有下列正式定义：

定义 6 如果对任意 $\varepsilon > 0$，都存在自然数 $N < \infty$，使得对任意 $n > N$，有

$$|x_n - x^*| < \varepsilon \tag{3.21}$$

那么称实数列 x_n 收敛于数 x^*，我们称 x^* 是 x_n 的极限．

37

简单地说，x_n 收敛于 x^* 意味着 x_n 可以在有限步后距离 x^* 任意近，现在可以提出两个重要的问题．

第一，如果 x_n 是随机变量而不是确定的值，我们能得到 x_n 的极限吗？这个问题很有意义，因为随机变量 x_n 容易出现极端值，当 $n > N$ 时，可能突然偏离 x^* 很远．

第二，既然我们能用不同方式定义“靠近”，我们应该也可以用不同方式定义收敛，这些定义等价吗？

我们以后会回答这些问题．极限在无法通过直接计算得到而需要近似计算时非常重要．例如，我们可能会用数列的极限来定义积分．

3.4.1　导数

导数[⊖]的概念至少能从两种角度理解．第一，导数可以衡量函数的光滑性，可以衡量变量的变化率．例如，如果资产价格的轨迹非常不规则，其关于时间的导数可能不存在．

第二，导数可以描述一个变量对另一个变量的变化作何种反应．例如，已知标的资产

⊖ 读者不要把数学中的导数(derivative)和金融中的衍生品(derivative securities)两个词弄混．

价格发生变化，我们可能想知道其期权价格将作何改变．这种导数通常可以通过链式法则求出．

导数是一种变化率，但它针对的是无穷小的变动．下面我们给出严格的数学定义．

定义7 令

$$y = f(x) \tag{3.22}$$

是定义域为 $\mathbf{R}$ 的函数．如果

$$f_x = \lim_{\Delta \to 0} \frac{f(x+\Delta) - f(x)}{\Delta} \tag{3.23}$$

存在，那么称 $f(x)$ 关于 x 的导数存在，记作 f_x，其中 Δ 是 x 的变化量．

变量 x 可以和任何现实生活现象结合起来．假如 x 代表时间⊖，那么 Δ 代表一段有限的时间，$f(x)$ 代表 x 时刻 y 的值．$f(x+\Delta)$ 代表 $x+\Delta$ 时刻 y 的值．因此，(3.23)中分子的值就是时间段 Δ 内 y 的变化量．(3.23)表示同样时间间隔内 y 的变化率．例如，如果 y 表示当时间为 x 时某种资产的价格，那么(3.23)表示时间间隔 Δ 内价格变化率．

为什么(3.23)中要用到极限呢？在定义导数时，极限有实际的作用，其作用是使得最后的比率与时间段 Δ 的大小无关．

要使比率与 Δ 的大小无关，需要付出代价．导数是基于无穷小时间段定义的，如果用较大的时间间隔来近似计算，其精度会下降．

1. 例子：指数函数

用指数函数来作为讨论导数的例子，设

$$f(x) = Ae^{rx}, \quad x \in \mathbf{R} \tag{3.24}$$

图3-3是当 $r>0$ 时该函数的图像．将其关于 x 求导：

38

$$f_x = \frac{\mathrm{d}f(x)}{\mathrm{d}x} = r[Ae^{rx}] = rf(x) \tag{3.25}$$

f_x 的值是 $f(x)$ 在 x 处的变化率．注意到当 x 增大时，e^{rx} 也增加．从图3-3可以看出 $f(\cdot)$ 的增速逐渐加快．比率

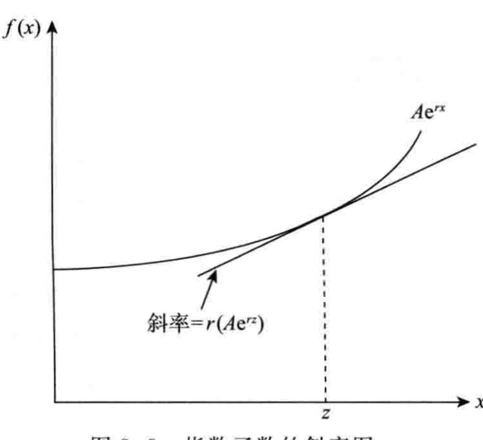


图3-3 指数函数的斜率图

$$\frac{f_x}{f(x)} = r \tag{3.26}$$

是变化的百分率．我们可以发现指数函数关于 x 变化的百分率是一个常数．

2. 用导数进行近似计算

为了研究如何利用导数进行近似计算，考虑下列情形．

令 Δ 为一有限间隔，根据式(3.23)导数的定义，若 Δ 较小，则有下列近似：

$$f(x+\Delta) \approx f(x) + f_x \Delta \tag{3.27}$$

⊖ 时间是我们能想到的为数不多的确定性变量之一．

39

该式表明 $f(\cdot)$在 $x+\Delta$ 的值可通过 $f(\cdot)$在 x 的值加上导数 f_x 与 Δ 的积近似得出．因此，当你不知道 $f(x+\Delta)$是多少的时候，已知 $f(x)$、f_x 和 Δ 就足以得到一个近似值．[⊖]

图 3-4 为该过程的示意图，其中比率

$$\frac{f(x+\Delta)-f(x)}{\Delta} \tag{3.28}$$

为线段 AB 的斜率．A 点固定，随着 Δ 逐渐减小，线段 AB 收敛到点 A 处的切线，因此，f_x 就是 A 点处的切线斜率．

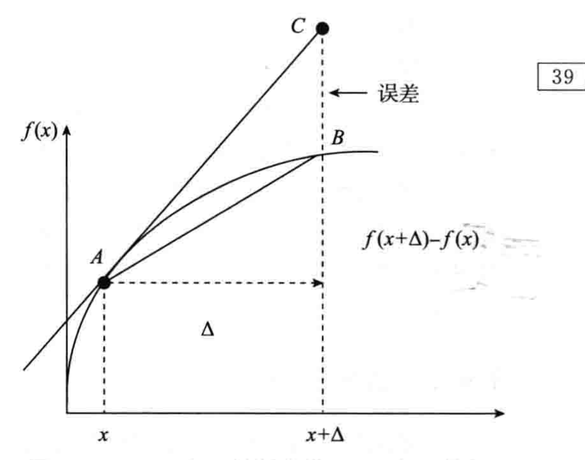


图 3-4　$f(x)$在 x 的导数是 $f(x)$在 x 的切线斜率的示意图

当我们将 $f_x\Delta$ 与 $f(x)$相加，得到 C 点的值，C 点的值可看作 B 点的估计值，其准确度取决于 Δ 的大小和函数 $f(\cdot)$的形状．

这一点可以用两个简单的例子来说明．第一，考虑图 3-5，此时 Δ 较大，不出所料，近似值$f(x)+f_x\Delta$ 距离 $f(x+\Delta)$不是很近．

图 3-6 描述了一个更有意义的例子．我们考虑一个不是很光滑的函数 $f(\cdot)$，同样用

$$\hat{f}(x+\Delta)=f(x)+f_x\Delta \tag{3.29}$$

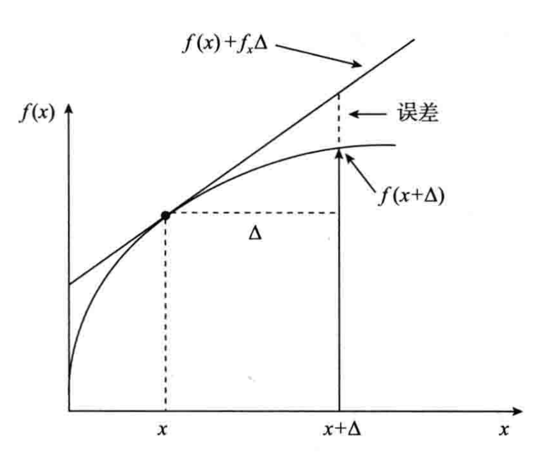


图 3-5　用 $f(x)$加上 $f(x)$在 x 点的导数乘以 Δ 作为 $f(x+\Delta)$的近似示意图

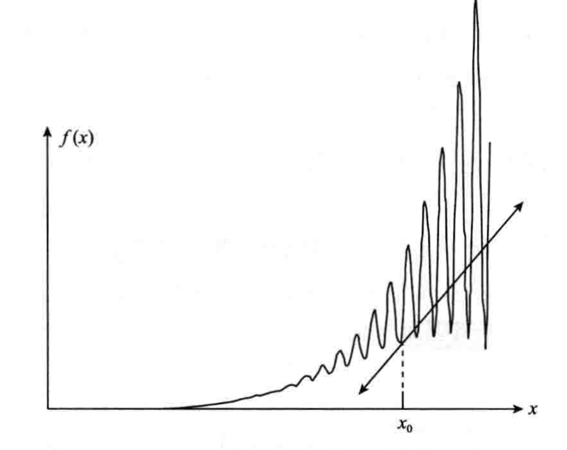


图 3-6　一个不是很光滑的函数的斜率示意图

得到估计值 $\hat{f}(x+\Delta)$，它的近似效果非常糟糕．显然，函数 $f(\cdot)$波动越剧烈，这种近似方法效果越差．

⊖ 如果 x 代表现在，那么 $f(x+\Delta)$属于未来．但是，$f(x)$、f_x 和 Δ 都是和"现在"相关的量．在这个意义下，实际中我们可以用这些量粗略地预测 $f(t+\Delta)$．该预测需要知道函数在 x 点导数 f_x 的数值．

下面考虑一个极端的例子．

3. 例子：高变差的情况

观察图 3-7，其中函数 $f(x)$是连续的，但在很小的区间 Δ 内，其波动频率非常高．此时，不仅用

40

$$f(x+\Delta) \approx f(x)+f_x\Delta \tag{3.30}$$

来估计会失败，而且可能连 f_x 都不存在．例如，x_0 点函数 $f(x)$的变化率应该取多少呢？很难说．事实上，你能以 A 为切点画出很多斜率不同的切线，这意味着 $f(x)$在 A 点是不可导的．

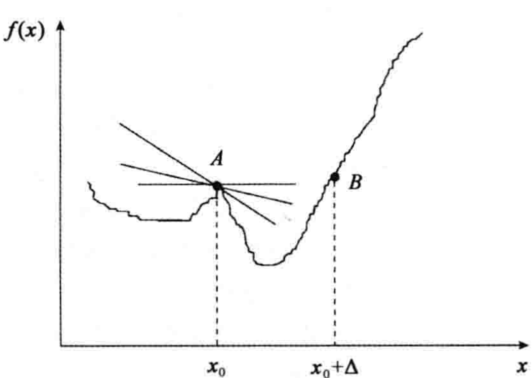


图 3-7　一个连续函数在一个小区间上有极端变差的示意图

3.4.2　链式法则

求导的第二个应用是链式法则．在之前讨论的例子中，$f(x)$是 x 的函数，x 代表时间，导数被解释为时间的改变对另一个变量的影响．

在衍生证券定价中，我们面对的问题有所变化．衍生资产(比如看涨期权)的价格依赖于标的资产的价格，而标的资产的价格又依赖于时间．[⊖]

因此，存在链式效应．随着时间推移，发生新(小)事件，标的资产的价格改变，衍生资产的价格受到影响．在常规微积分中，链式法则可以用于研究此类链式效应．

假如 $f(x)$中的 x 不直接表示时间，而是一个关于时间 $t(t\geqslant 0)$的函数：

$$x_t = g(t) \tag{3.31}$$

那么 $f(\cdot)$是一个复合函数，可表示为：

$$y_t = f(g(t)) \tag{3.32}$$

接下来的问题是如何用公式来表示 t 的变化对 y_t 的影响．

在常规微积分中链式法则定义如下：

定义 8 对于上述的 f 和 g，我们有

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{\mathrm{d}f(g(t))}{\mathrm{d}g(t)}\frac{\mathrm{d}g(t)}{\mathrm{d}t} \tag{3.33}$$

可以看到，链式法则是将两个导数相乘．首先，$f(g(t))$关于 $g(t)$求导．然后，$g(t)$关于 t 求导，最后将两个导数相乘，所得的积代表了 t 对 y_t 的影响．

链式法则在评估一个变量对另一个变量的影响时非常有用．

以衍生资产的价格为例．一名交易员已知标的资产价格的连续变化情况，并且希望知道关于该资产的复杂衍生品价值将如何变化．如果这些衍生品是在市场上流通的，其价格

⊖ 随着时间流逝，合约的到期日临近，即使标的资产价格保持不变，其看涨期权价格也会下降．

变化可以从市场上直接得到.㊀但如果该衍生品还未在市场上交易，其价格需要利用定价模型来计算. 这些定价模型会用到(3.33)中的链式法则.

在刚才的例子中，$f(x)$是x_t的函数，而x_t是个确定性变量，x_t中不包含随机成分. 假如x_t是随机的，或者$f(\cdot)$还依赖于某个随机变量z_t，又会怎样呢？也就是说：

41

1）我们还能使用原来的链式法则吗？

2）链式法则公式在随机环境下会发生怎样的变化？

第一个问题的答案是不能.(3.33)中的链式法则不能在连续时间的随机环境下使用. 后面我们将通过随机微积分来解决这一问题.

随机微积分与常规微积分目的相同，只是规则不同而已.

3.4.3 积分

积分是一种用于求和的数学工具. 运算符$\sum$是用于求可数个项的和，积分则是用于求不可数的无穷多个项的和. 由于求不可数个项的和比较令人费解，我们将给出积分的正式定义.

通常我们用一种直观的方法来理解积分. 我们可以从可数个项开始，得到一个近似值，然后利用极限推广到不可数个项. 使用不同的极限定义，就能够得到不同的积分. 最常见的积分是黎曼(Riemann)积分，还有一种类似方式定义的更一般的积分，叫做黎曼-斯蒂尔切斯(Riemann-Stieltjes)积分.

本节我们将回顾这些定义.

1. 黎曼积分

已知$f(t)$是$t\in[0,T]$的确定性函数. 假如我们想知道该函数在$[0,T]$的积分

$$\int_0^T f(s)\mathrm{d}s \tag{3.34}$$

参考图3-8.

42

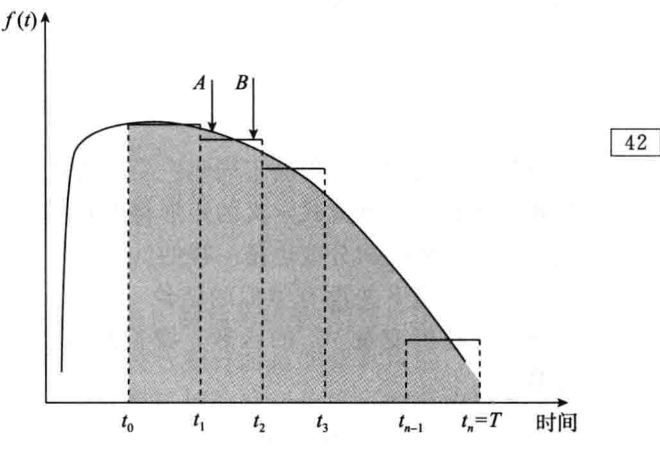


图3-8 用黎曼积分求曲线下方图形面积的构造步骤

要计算黎曼积分，我们首先将区间$[0,T]$分为n个互不重叠的子区间

$$0=t_0<t_1<\cdots<t_n=T \tag{3.35}$$

则积分的值可以近似为

$$\sum_{i=1}^{n} f\left(\frac{t_i+t_{i-1}}{2}\right)(t_i-t_{i-1}) \tag{3.36}$$

定义9 令$\max|t_i-t_{i-1}|\to 0$，称下列极限

$$\sum_{i=1}^{n} f\left(\frac{t_i+t_{i-1}}{2}\right)(t_i-t_{i-1})\to\int_0^T f(s)\mathrm{d}s \tag{3.37}$$

㊀ 当然，那个时候资产是否被正确定价是一个问题.

为 $f(t)$在$[0,T]$上的黎曼积分，其中左边是将 n 个矩形的面积相加，$|t_i-t_{i-1}|$为矩形的底边，$f(|t_i-t_{i-1}|/2)$为矩形的高，图 3-8 展示了这种结构．注意，A 的面积近似等于 B 的面积，这在矩形底边很短或者 $f(t)$较平缓（不在小区间内大幅波动）时尤为明显．

假如矩形的面积之和不能够估计曲线下的面积，我们可以尝试用另一种分割来达到目的．随着$|t_i-t_{i-1}|$不断缩小，矩形的底边不断缩短，矩形的数量增加，面积的估计值将更加准确．

显然，要用这种方法估计，$f(t)$较平缓的条件是不可缺少的．事实上，如果 $f(t)$的形状非常不规则，这种方法将难以进行．用专业术语来说，这种方法要行得通，函数必须是黎曼可积的．

图 3-9 给出了一个反例，在这里 $f(t)$波动剧烈．如果在矩形缩小后这样的波动没有变得可控，这种用矩形估计的方法很可能失败．

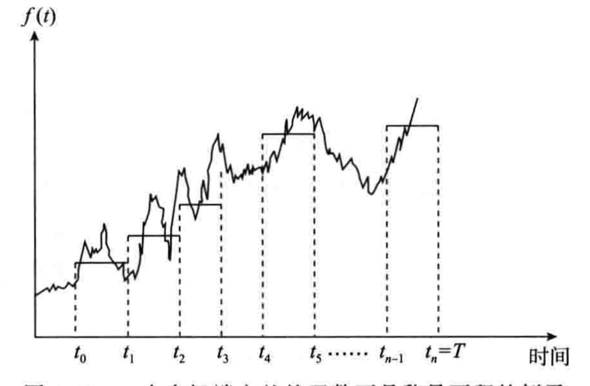


图 3-9　一个有极端变差的函数不是黎曼可积的例子

43

此外，还有一点在后面伊藤积分的学习中非常重要．用于估计曲线下面积的矩形是通过特殊规则构造的．我们是用 $f(t)$在区间$[t_{i-1},t_i]$中点的值作为矩形的高，那么如果用其他方法定义矩形，该估计方法还可行吗？例如，我们用

$$f(t_i)(t_i-t_{i-1}) \tag{3.38}$$

或

$$f(t_{i-1})(t_i-t_{i-1}) \tag{3.39}$$

来计算矩形面积，结果会不同吗？要回答这个问题，考虑图 3-10，注意到随着分割越来越细化，用以上三种方式定义的矩形将得到同一个估计值．因此，即使定义矩形时使用的高度不同，得到的积分值也是一样的．

随机环境下并没有类似的结论．假定 $f(W_t)$是随机变量 W_t 的函数，我们想要计算

$$\int_0^T f(W_s)\mathrm{d}W_s \tag{3.40}$$

与确定性的问题不同，取矩形面积为

$$f(W_{t_i})(W_{t_i}-W_{t_{i-1}}) \tag{3.41}$$

和取为

$$f(W_{t_{i-1}})(W_{t_i}-W_{t_{i-1}}) \tag{3.42}$$

将得到不同的结果．

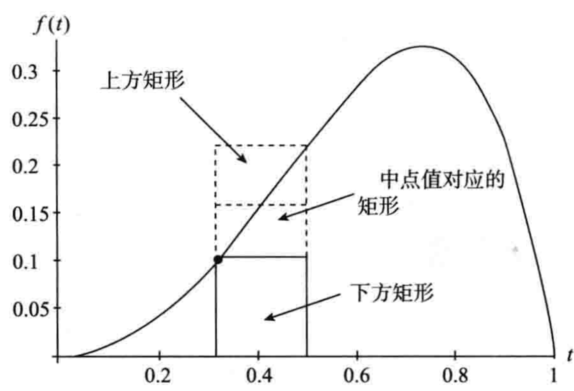


图 3-10　分别使用上方矩形、下方矩形、中点值对应的矩形近似曲线下方图形面积示意图

要理解其背后的原因，考虑 W_t 是鞅

的情况．此时(3.42)中，以t_{i-1}时刻的信息为条件得到的$(W_i - W_{i-1})$条件期望会等于零．这是因为根据定义，鞅的未来增量与当前信息集无关．

另一方面，(3.41)中的条件期望项一般是非零的．[⊖]所以，在随机微积分中，用不同方法定义矩形会产生不同的结果．

最后，我们强调一下这一重要结论：当$f(\cdot)$是随机变量的函数时，其积分也是随机变量，因此必须使用随机积分．

2. 斯蒂尔切斯积分

斯蒂尔切斯积分是另一种定义的积分．定义微分 $\mathrm{d}f$ 是 x 的无限小波动引起的函数值$f(x)$的变化：

$$\mathrm{d}f(x) = f(x+\mathrm{d}x) - f(x) \tag{3.43}$$

此外我们已介绍过等式：

$$\mathrm{d}f(x) = f_x(x)\mathrm{d}x \tag{3.44}$$

(注意到在上述标记中，$f_x(x)$也是 x 的函数．)假如我们要求 $h(x)$关于 x 的积分：

44

$$\int_{x_0}^{x_n} h(x)\mathrm{d}x \tag{3.45}$$

其中 $h(x)$为

$$h(x) = g(x)f_x(x) \tag{3.46}$$

那么斯蒂尔切斯积分可定义为

$$\int_{x_0}^{x_n} g(x)\mathrm{d}f(x) \tag{3.47}$$

其中

$$\mathrm{d}f(x) = f_x(x)\mathrm{d}x \tag{3.48}$$

该定义与黎曼积分区别不大．可以说两者采用相似的原理．

若 x 代表时间 t，在分割的区间$[0,T]$上，斯蒂尔切斯积分可写为

$$\int_0^T g(s)\mathrm{d}f(s) \approx \sum_{i=1}^n g\left(\frac{t_i + t_{i-1}}{2}\right)(f(t_i) - f(t_{i-1})) \tag{3.49}$$

由于这些相似性，右端项在 $\max_i |t_i - t_{i-1}|$ 趋于 0 时的极限被称为黎曼-斯蒂尔切斯积分．

黎曼-斯蒂尔切斯积分在关于 $f(x)$而非 x 的增量进行积分时是非常有用的，而在处理金融衍生品时，通常是这种情况．衍生资产的价格依赖于标的资产价格，而标的资产价格是时间的函数．因此黎曼-斯蒂尔切斯积分比黎曼积分更适用于处理衍生资产价格．

然而，在得到该结论前，必须意识到所有讨论都是基于函数关于时间是在确定性关系这一基础之上．在随机环境中该定义还有效吗？我们还能用矩形估计积分吗？改变定义矩形的方式会对结果产生影响吗？

总的来说，这些问题的答案是否定的．随机环境下函数积分无法由黎曼积分直接推广

⊖ 注意$(W_{t_i} - W_{t_{i-1}})$和 W_{t_i} 是相关的．

得到．因此，我们需要一种新的方式定义积分．

3. 例子

本节我们将选取一个简单的函数来作为讨论黎曼一斯蒂尔切斯积分的例子．令

$$g(S_t)=aS_t \tag{3.50}$$

其中 a 是常数，$g(\cdot)$是 S_t 的线性函数．[⊖]在黎曼-斯蒂尔切斯定义下，积分

$$\int_0^T aS_t\mathrm{d}S(t) \tag{3.51}$$

的值是多少呢？

直接运算，得

$$\int_0^T aS_t\mathrm{d}S(t)=a\left[\frac{1}{2}S_t^2\right]_0^T \tag{3.52}$$

或

$$\int_0^T aS_t\mathrm{d}S_t=a\left[\frac{1}{2}S_T^2-\frac{1}{2}S_0^2\right] \tag{3.53}$$

现在，看看能否通过矩形面积的近似方法得到同样的结果．

由于 $g(\cdot)$是线性的，该估计方法效果很好．我们用底边中点的函数值作为矩形的高恰好是准确的．图 3-11 描述了该过程，其中 $a=4$.

由于 $g(\cdot)$是线性函数，以区间中点的函数值作为高得到的单个矩形刚好能够代替阴影部分的面积．事实上，矩形 S_0ABS_T 的面积是

$$a\left[\frac{S_T+S_0}{2}\right][S_T-S_0]=a\left[\frac{1}{2}S_T^2-\frac{1}{2}S_0^2\right] \tag{3.54}$$

黎曼-斯蒂尔切斯方法刚好能够测量该矩形面积，无论用于估测的小矩形数量是多少．

45

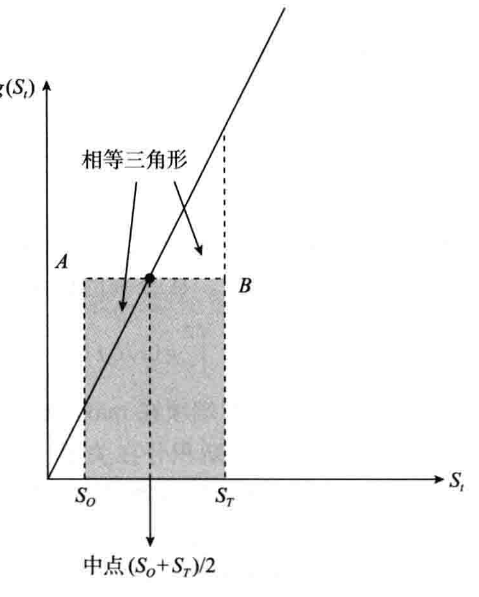


图 3-11 一个用矩形很好地近似的例子

3.4.4 分部积分

在常规微积分中，分部积分是一个重要的结论．它能够将一些积分变形成容易处理的形式．随机微积分中也有类似的公式，不过形式有所不同．

考虑两个可微函数 $f(t)$和 $h(t)$，其中 $t\in[0,T]$代表时间．有以下关系：

$$\int_0^T f_t(t)h(t)\mathrm{d}t=[f(T)h(T)-f(0)h(0)]$$

⊖ S_t 是时间的函数．

$$-\int_0^T f(t)h_t(t)\mathrm{d}t \tag{3.55}$$

其中 $h_t(t)$和 $f_t(t)$是对应的函数关于时间 t 的导数．它们也是 t 的函数．

结合斯蒂尔切斯积分的概念，该变形意味着能够将

$$\int_0^T h(t)\mathrm{d}f(t) \tag{3.56}$$

变形成含有

$$\int_0^T f(t)\mathrm{d}h(t) \tag{3.57}$$

的表达式．

该变形在随机环境下对计算伊藤积分非常有用．事实上，假定 $f(t)$是随机的但在一定条件下是时间 t 的确定性函数，那么通过分部积分，我们能够将随机积分写成关于一个确定性变量的积分函数．在随机微积分中，这一点将在伊藤公式里用到．

3.5 偏导数

对于一个看涨期权，到期日将从两个方面影响期权价格．首先，随着时间推移，到期日临近，期权剩余时间减少，这将拉低价格．但另一方面，随着时间推移，标的资产的价格发生变化，也会影响期权价格．因此，该看涨期权价格是两个变量的函数．可以写成

$$C_t = F(S_t, t) \tag{3.58}$$

其中 C_t 是看涨期权价格，S_t 是标的资产价格，t 是时间．

现我们将 t 固定，对 $F(S_t,t)$关于 S_t 微分，得到的偏导数是

$$\frac{\partial F(S_t,t)}{\partial S_t} = F_s \tag{3.59}$$

(理论上)上式表示当 t 固定时，标的资产价格变化对期权价格的影响．这种影响是抽象的，因为实际中，当 t 固定时，S_t 不会发生变化．

46

同样，对时间变量的偏导数定义为

$$\frac{\partial F(S_t,t)}{\partial t} = F_t \tag{3.60}$$

注意，尽管 S_t 是时间的函数，在这里我们假定它是不变的，同样，这仍然是一种抽象的假设，因为随着 t 的变化，S_t 也会变化．但求偏导时，我们把它当作常数．

由于偏导数的这种抽象特征，它不能直接用于解释金融市场上资产价格的实际变化．但是偏导数是一种非常有用的间接工具，它能将整个变化分割成几个来自不同因素的部分，而且它奠定了全微分的基础．

在介绍全微分前，关于偏导数还有一点需说明：由于偏导数不代表观察到的变化，应用它时，确定性和随机性环境中没有区别．因此，我们不必为偏导数在随机环境中作新的定义．

为了更好地明白这一点，请看以下例子．

3.5.1 例子

考虑有两个变量的函数

$$F(S_t,t)=0.3S_t+t^2 \tag{3.61}$$

其中 S_t 是某金融资产的(随机)价格，t 是时间．

对 F 关于 S_t 求偏导：

$$\frac{\partial F(S_t,t)}{\partial S_t}=0.3 \tag{3.62}$$

这里 ∂S_t 是抽象的 S_t 的增量，不等同于实际中 S_t 的增加．事实上偏导数 F_S 仅表示当 S_t 变化一个单位后 $F(\cdot)$将如何变化，F_S 只是一个全微分中的乘数．

3.5.2 全微分

假如我们在时刻 t 观察到看涨期权价格的一个小变化，记为 $\mathrm{d}C_t$. 该变化中有多少比例是标的资产价格的变化造成的？有多少比例是由到期日的临近造成的？全微分可以解答这些问题．

假定 $f(S_t,t)$是双变量函数，它的全微分定义为：

$$\mathrm{d}f=\left[\frac{\partial f(S_t,t)}{\partial S_t}\right]\mathrm{d}S_t+\left[\frac{\partial f(S_t,t)}{\partial t}\right]\mathrm{d}t \tag{3.63}$$

也就是说，我们用 S_t 的变化量 $\mathrm{d}S_t$ 乘以偏导数 f_S，用时间的变化量 $\mathrm{d}t$ 乘以偏导数 f_t. $f(\cdot)$的总变化即为二者之和．由此可知，全微分是将一个观察到的完整变化分割成若干抽象部分进行计算．

3.5.3 泰勒展开式

1. 一元泰勒展开

假设 $f(x)$为无穷阶可微函数，$x\in\mathbf{R}$，任取 x 记为 x_0.

定义 10 $f(x)$在 $x_0\in\mathbf{R}$ 附近的泰勒展开式为：

$$f(x)=f(x_0)+f_x(x_0)(x-x_0)+\frac{1}{2}f_{xx}(x_0)(x-x_0)^2+\frac{1}{3!}f_{xxx}(x_0)(x-x_0)^3+\cdots$$

47

$$=\sum_{i=0}^{\infty}\frac{1}{i!}f^i(x_0)(x-x_0)^i \tag{3.64}$$

其中 $f^i(x)$是 $f(x)$关于 x 的 i 阶导数．$f^i(x_0)$为其在 x_0 点的取值．[⊖]

(3.64)为什么在 $f(x)$为连续且足够光滑时成立，我们不做详述．泰勒展开式的正确性毋庸置疑，我们仅讨论它的一些含义．

首先注意到(3.64)的表达式中不是近似等号，而是等号，等式右端包含了无穷多项，

⊖ 这意味着 x_0 代入 $f^i(\cdot)$后，该值将成为与 x 无关的常数．

每一项都包含 x 的幂次项，但这样的元素有无穷多个．正因为这一点，泰勒展开式在实践中的实用性受到了限制．

但是，我们可以用展开式(3.64)来得到近似值．考虑等式(3.64)，仅取 x_0 附近的 x，即假定

$$(x-x_0)\approx\text{“很小”} \tag{3.65}$$

那么必然有

$$|x_1-x_0|>|x_1-x_0|^2>|x_1-x_0|^3>\cdots \tag{3.66}$$

($|x_1-x_0|<1$，则幂次越高，值越小．)

在这些条件下，我们可以考虑舍弃(3.64)右边的一些项，只要我们能够论证它们是可忽略的．因而我们要约定精确程度，然后以此为标准对各项进行筛选．但要多小才可以忽略呢？

微积分中的约定通常是，如果 x 是确定性变量，$(\mathrm{d}x)^2$ 或更高阶的项是可以忽略的．[⊖]因此，只要我们假定 x 是确定性的，且$(x-x_0)$较小，那么我们就可以使用一阶泰勒公式来估计：

$$f(x)\approx f(x_0)+f_x(x_0)(x-x_0) \tag{3.67}$$

假如 $f(x)$在 x_0 一阶可导，该不等式就可以成为一个等式，只要

$$(x-x_0)\to 0 \tag{3.68}$$

在该前提下，无穷小的偏离$(x-x_0)$可以记作

$$\mathrm{d}x\approx(x-x_0) \tag{3.69}$$

同样，$f(\cdot)$的无穷小偏离可以表示为

$$\mathrm{d}f(x)\approx f(x)-f(x_0) \tag{3.70}$$

因此，我们可以得到微分中熟悉的结论：

$$\mathrm{d}f(x)=f_x(x)\mathrm{d}x \tag{3.71}$$

这里 $f_x(x)$是 x 的函数而不是 $f_x(x_0)$，因为我们考虑的是当 x 趋向 x_0 时的极限．

2. 二阶估计方法

等式

$$f(x)\approx f(x_0)+f_x(x_0)(x-x_0) \tag{3.72}$$

叫做一阶泰勒估计式．通常，如果将二次项加进来，可以得到更精确的估计值：

$$f(x)\approx f(x_0)+f_x(x_0)(x-x_0)+\frac{1}{2}f_{xx}(x_0)(x-x_0)^2 \tag{3.73}$$

这一点与后面介绍的随机微积分的内容关系密切．为了更好地引入伊藤引理，我们来看一个具体例子．

3. 久期和凸性

考虑以下指数函数，其中 t 代表时间，T 为固定值，$r>0$，$t\in[0,T]$：

⊖ 如果这样，$(\mathrm{d}x)^3$，$(\mathrm{d}x)^4$，…，将比$(\mathrm{d}x)^2$ 更小．

$$B_t = 100\mathrm{e}^{-r(T-t)} \tag{3.74}$$

48

该函数在 $t=0$ 处有初始值 $B_0=100\mathrm{e}^{-rT}$，然后以 $r\%$的增长率增加．当 $t\to T$ 时，B_t 的值趋向 100，因此可以将 B_t 视为 t 时刻的价值，T 时的价值为 100. 它可以看作一份 T 时刻到期的无违约风险零息债券的现价．r 是以连续复利计算的利率．

我们希望用关于 t 的泰勒公式估计 B_t. 假定 r，T 是常数，$t=t_0$ 附近的一阶泰勒展开为

$$B_t \approx 100\mathrm{e}^{-r(T-t_0)} + (r)100\mathrm{e}^{r(T-t_0)}(t-t_0), t\in[0,T] \tag{3.75}$$

其中右端第一项是 B_t 在 t_0 的值，第二项是 B_t 一阶导在 t_0 的值乘以时间增量 $t-t_0$.

图 3-12 是这种近似思想的示意图，B_t 由上凹曲线表示，一阶泰勒展开式由直线表示，两者相切于 A 点．可以看到，t 无论从哪个方向远离 t_0，估计值的精确度都会大幅下降，但在 t_0 附近，近似值的精确度很高．

图 3-13 给出了用二阶泰勒公式进行近似的情况：

$$\begin{aligned}B_t \approx & 100\mathrm{e}^{-r(T-t_0)} + (r)100\mathrm{e}^{-r(T-t_0)}(t-t_0) \\ & + \frac{1}{2}(r^2)100\mathrm{e}^{-r(T-t_0)}(t-t_0)^2, \quad t\in[0,T]\end{aligned} \tag{3.76}$$

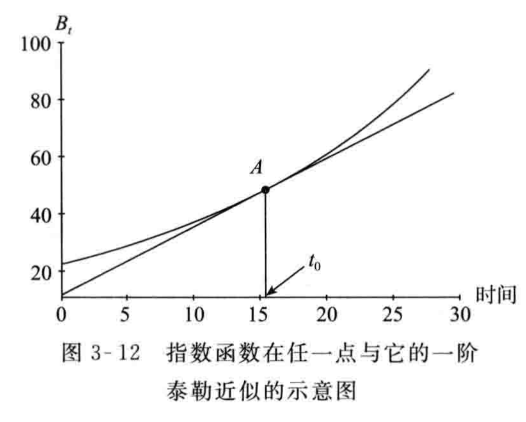


图 3-12　指数函数在任一点与它的一阶泰勒近似的示意图

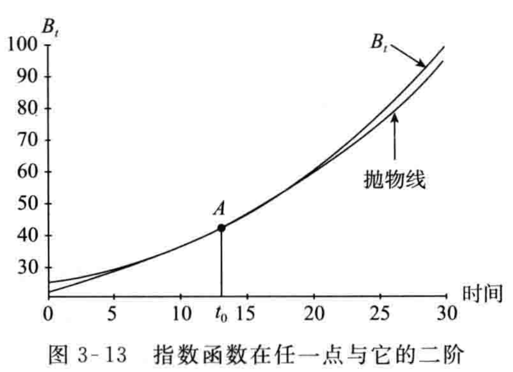


图 3-13　指数函数在任一点与它的二阶泰勒近似的示意图

等式右端对应的图像是与指数曲线在 A 点相切的抛物线．由抛物线在 A 点的凸性，可以推断该方法的近似效果比一阶泰勒近似更好．

一阶和二阶泰勒公式的差异在于$(t-t_0)^2$ 项．t 越接近 t_0，该项越小．而且，该项比$(t-t_0)$项更快地趋于 0.

以上泰勒估计式反映了债券的折现价格是如何随着时间的变化而变化的．

另一种泰勒估计式是固定 t 和 T，将 B_t 关于 r 展开得到的．考虑 r_0 附近的二阶展开式：

49

$$B_t \approx [100\mathrm{e}^{-r_0(T-t)}]\left[1-(T-t)(r-r_0)+\frac{1}{2}(T-t)^2(r-r_0)^2\right], \quad t\in[0,T], \quad r>0$$

或者，除以 $100\mathrm{e}^{-r_0(T-t)}$

$$\frac{\mathrm{d}B_t}{B_t} \approx -(T-t)(r-r_0)+\frac{1}{2}(T-t)^2(r-r_0)^2$$

以上二阶泰勒展开给出了 r 发生无穷小变化后该零息债券价值变化的百分率．右端项表示 r 变化 $r-r_0$ 后债券价值的变化率．其中 r_0 是当前利率．右端有 2 项包含$(r-r_0)$，在金融市场中第一项的系数叫做修正久期．第二项为正且系数为$\frac{1}{2}(T-t)^2$，该项即所谓的债券的凸性．总而言之，B_t 关于 r 的二阶泰勒展开式说明债券价值随利率的上升(下降)而下降(上升)；凸性使得利率下降或上升同样的值时，债券价格的上升的幅度大于价格下降的幅度．

4. 多元泰勒展开公式

$f(x)$是无穷阶可微函数，$x\in\mathbf{R}^d$，其中 $x=(x_1,\cdots,x_d)$，任取 x，记作 a.

定义 11 $f(x)$在 $a\in\mathbf{R}^d$ 附近的泰勒展开式定义为

$$f(x)=f(x_1,\cdots,x_d) \tag{3.77}$$

$$=f(a_1,\cdots,a_d)+\sum_{i=1}^{d}\frac{\partial f(x)}{\partial x_i}(x_i-a_i) \tag{3.78}$$

$$+\frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\frac{\partial^2 f(x)}{\partial x_i\,\partial x_j}(x_i-a_i)(x_j-a_j) \tag{3.79}$$

$$+\frac{1}{3!}\,\frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{k=1}^{d}\frac{\partial^3 f(x)}{\partial x_i\,\partial x_j\,\partial x_k}(x_i-a_i)(x_j-a_j)(x_k-a_k)+\cdots \tag{3.80}$$

随机波动模型中的偏微分方程常出现混合偏导数，要逼近这些混合偏导数，我们要用到高阶泰勒展开式．

3.5.4 常微分方程

我们要回顾的第三个概念是常微分方程．例如，考虑以下方程：

$$\mathrm{d}B_t=-r_tB_t\mathrm{d}t \tag{3.81}$$

其中 B_0 已知，$r_t>0$. 该式说明 B_t 是受 t 影响的量，例如 B_t 的变化量是 t 和 B_t 的函数，该方程称为常微分方程．B_t 变化的百分比等于 r_t 与 $\mathrm{d}t$ 的乘积：

$$\frac{\mathrm{d}B_t}{B_t}=-r_t\mathrm{d}t \tag{3.82}$$

现定义 B_t：

$$B_t=\mathrm{e}^{-\int_0^t r_u\mathrm{d}u} \tag{3.83}$$

将其带入(3.81)，它满足方程，故(3.83)是该常微分方程的一个解．

常微分方程首先是一个方程，它是一个含有未知函数及其导数的方程．

可以用一个简单的例子来理解，在方程

$$3x+1=x \tag{3.84}$$

中，x 是未知的，有待求解．这里方程的解是 $x=-\frac{1}{2}$.

在矩阵方程

$$Ax - b = 0 \tag{3.85}$$

50
未知元素 x 是一个向量．在一定条件下，该方程的解为 $x=A^{-1}b$，即 A 的逆矩阵乘以 b.

在常微分方程

$$\frac{\mathrm{d}x_t}{\mathrm{d}t} = ax_t + b \tag{3.86}$$

中，x_t 未知，且 x_t 是一个函数．更准确地说，它是 t 的函数：

$$x_t = f(t)$$

再来看本节提到的常微分方程：

$$\mathrm{d}B_t = - rB_t \mathrm{d}t \tag{3.87}$$

在 $B_T=1$ 的条件下，解为：

$$B_t = \mathrm{e}^{-\int_0^t r_u \mathrm{d}u} \tag{3.88}$$

该函数其实是零息债券的价值函数．这个例子说明固定收益证券的定价公式可能是某个微分方程的解．在随机环境下，我们还会遇到很多这样的例子．

最后，我们定义积分方程

$$\int_0^t (ax_S + b)\mathrm{d}S = x_t \tag{3.89}$$

其中未知量 x_t 是 t 的函数．

3.6 小结

本章回顾了微积分的基本知识，大多数概念是基础性的．尽管微分、积分、泰勒公式等概念并不陌生，复习这些内容对以后的学习很有帮助．

随机微积分的内容与之相似，但所研究的现象是连续时间下的随机过程．在随机环境下，常规定义的微分、积分、泰勒展开都不适用．要理解随机微积分中的概念，首先必须深刻理解对应的常规概念．

本章另一个重要概念是无穷小量，尤其是我们需要约定小到什么程度的增量是可以忽略的.

3.7 参考阅读

现在你可以阅读一些初等微积分教材．复习一下基础的求导、积分法则，做一些练习题，都能够帮助你掌握这些内容．

3.8 习题

1. 对于数列 $\{X_n\}$，$n=1, 2, 3, \cdots$，如果

(a) $X_n = a_n$

(b) $X_n = \left(1+\frac{1}{n}\right)^n$

(c) $X_n = (-1)^{n-1}/n$

(d) 以上数列是否收敛？

(e) 假设年利率是5%，Δ是时间间隔，一年有 n 个 Δ，即 $n\Delta=1$.

ⅰ. 1美元的投资经过Δ后，总收入是多少？

ⅱ. 现假设期限为Δ的短期国债，年利率为5%，一年的复合回报是多少？

2. 对于 $n=1, 2, 3, \cdots$，求以下数列的极限(如果存在)：

(a) $x_n=(-1)^n$

(b) $x_n=\sin\left(\frac{n\pi}{3}\right)$

(c) $x_n=n(-1)^n$

(d) $x_n=\sin\left(\frac{n\pi}{3}\right)+(-1)^n/n$. 该数列是否有界？

51

3. 求下列极限：

$\lim\limits_{n\to\infty}(3+\sqrt{n})/\sqrt{n}$

$\lim\limits_{n\to\infty}n^{1/n}$

4. 证明部分和

$$S_n=\sum_{k=1}^{n}\frac{1}{k!}$$

当 $n\to\infty$ 时是收敛的.

5. 证明按以下方式定义的部分和收敛到3(使用数学归纳法)，其中 $S_1=1$：

$$S_{n+1}=\sqrt{3S_n}$$

6. 当 $N\to\infty$ 时，级数 $\sum\limits_{n=1}^{N}\frac{1}{n}$ 是否收敛？

7. 假定

$$X_n=aX_{n-1}+1$$

且 X_0 已知. 在什么条件下，X_n 的部分和收敛？

8. 考虑函数

$$f(x)=x^3$$

(a) 计算积分

$$\int_0^1 f(x)\mathrm{d}x$$

(b) 将区间[0,1]分为4份(可以不均匀)以选择 x_i，计算以下求和项：

$$\sum_{i=1}^{4}f(x_i)(x_i-x_{i-1})$$

$$\sum_{i=1}^{4}f(x_{i-1})(x_i-x_{i-1})$$

(c) 以上两项相差了多少，哪一个更接近积分真实值？

9. 考虑以下函数：

$$f(x)=\begin{cases}x\sin\left(\frac{\pi}{x}\right) & 0<x<1\\ 0 & x=0\end{cases}$$

(a) 计算积分 $\int_0^1 f(x)\mathrm{d}x$.

(b) 将区间[0, 1]分为4份，$0=x_0<x_1<x_2<x_3<x_4=1$，选择 x_i，并计算：

$$\sum_{i=1}^{4}f(x_i)(x_i-x_{i-1})$$

$$\sum_{i=1}^{4} f(x_{i-1})(x_i - x_{i-1})$$

(c) 以上求和项是否能近似真实积分？

(d) 为什么？

10. 考虑以下函数：

$$f(x,\ y,\ z)=\frac{x+y+z}{(1+x)(1+y)(1+z)}$$

分别计算关于 x，y，z 的偏导数.

52 ~ 53

11. 级数 $\sum_{n=1}^{N} \frac{1}{n^2}$ 在 $N \to \infty$ 时是否收敛？

12. 编写程序，用数值方法计算以积分 $\int_1^2 x^4 e^{2x} dx$.

第 4 章　衍生品定价：模型和记号

4.1　引言

衍生品的定价工具和普通资产的定价理论有着很大的区别．在一定的简化假定条件下，人们可以将衍生品的无套利价格写成一些“标的”证券的函数，然后得到一系列公式．建立这些公式不必考虑与其他金融市场的联系或者实物层面的经济因素．

有一些具体方法可以得到这样的公式．第 2 章讨论过一种方法，我们用套利的概念找到一种概率测度，使得金融资产折现后在该测度下满足鞅的条件，有鞅作为工具，只要得到新测度下的期望价值，我们就能轻松地算出无套利价格．这种衍生品定价方法称为等价鞅测度法．

另一种利用套利的定价方法在某种程度上更为直接，定价者首先构造一个无风险投资组合,然后以无套利为条件得到一个偏微分方程．该方程既有可能用解析法求解也有可能用数值法求解．

在之前的例子中，衍生品定价关键是要找到函数 $F(S_t, t)$，它将衍生品的价格和标的资产的价格以及一些其他市场风险因素联系起来．如果无法得到封闭形式的公式，也可以通过数值方法近似地描述这种关系．

本章针对如何得到线性或非线性衍生品的定价公式举了一些例子．我们阐明了相关概念并给出相应的偏微分方程方法．这方面的介绍将使我们更清楚为何稍后要介绍随机微积分的基本工具．

4.2　定价函数

衍生品定价的关键问题是得到函数 $F(S_t, t)$，其中 S_t 是标的资产的价格，t 是时间．一般情况下，人们会首先尝试得出 $F(S_t, t)$的封闭形式，Black-Scholes 公式是最著名的例子，它给出了看涨期权价格与标的资产以及其他一些相关参数的关系．事实上，类似的公式还有很多，有些更为简单．

当封闭形式的公式不存在时，人们会寻找包含 $F(S_t, t)$的某个方程．⊖

本节我们将举例说明如何得到 $F(S_t, t)$. 顺便我们将介绍新的数学工具和概念，这些工具和概念在衍生品的定价中非常有用．

⊖ 封闭形式的公式不存在并不意味着定价函数不存在．可能仅仅意味着不能用简单的方式表达定价函数．例如，所有连续和“光滑”的函数都可以改写成无穷项的泰勒展开式．同时，为得到封闭形式公式而舍去泰勒公式的项会产生误差．

4.2.1 远期合约

考虑一类现购自运货物[1]. 我们将介绍如何得到远期合约的定价公式，其中 S_t 是标的资产的价格. 假定远期合约包含以下条款：

- 在某未来的时刻 T，其中
$$t < T \tag{4.1}$$
一方将以每单位 F 美元的价格向另一方购买黄金.
- 合约在 t 时刻签订，在 T 之前不执行合约.

我们的合约对双方都施加义务，一方提供黄金，另一方进行购买. 根据已有条件，我们该如何得到该合约在 t 时刻的价格函数 $F(S_t,t)$呢[2]？我们要利用套利来进行研究.

假定在 t 时刻我们以无风险连续复利率 r_t 借入 S_t 美元并用这些钱买入 1 单位黄金，r_t 在合约期间固定不变，单位时间的保险费、储存费为 c 美元且在 T 时刻支付. 在 $T-t$ 时间内持有的 1 单位黄金的总成本是
$$e^{r_t(T-t)}S_t + (T-t)c \tag{4.2}$$
其中第一项代表 T 时刻归还给银行的本息之和，第二项代表 T 时刻支付的保险和存储费用.

56

这是在 T 时刻获得 1 单位黄金的方法之一. 即借入所需资金，买入 1 单位黄金，持有它直到 T 时刻.

另一种方法是使用远期合约. 你只需签署一份 T 时刻买入 1 单位黄金的合约，所有支付在到期日进行.

这两种方法的结果是完全相同的，否则会存在套利机会[3]. 套利者只需签订两份合约，在一份合约中买入黄金，同时在另一份中以更高的价格卖出. 故可以得到以下数学公式：
$$F(S_t,t) = e^{r_t(T-t)}S_t + (T-t)c \tag{4.3}$$
这就是我们分析各种套利可能性并结合 S_t、t 等参数得到的远期合约的价格函数 $F(S_t, t)$.

在以上讨论中，$F(S_t,t)$和 t 是变量，在合约执行前会变化. 另外，c，r_t，T 是参数，在合约 $T-t$ 期间假定不变.

由(4.3)可知，$F(S_t,t)$关于 S_t 是线性的. 因此远期合约是线性产品. 以后我们会推导出 Black-Scholes 公式，它是看涨期权的定价公式，该公式得到的期权价格关于 S_t 是非线性的. 类似于期权的衍生品一般都是非线性产品.

边界条件

首先我们简单介绍一下什么是边界条件. 假设用严格的数学符号表达"到期日临近"的

---

[1] 参照第 1 章中的定义.

[2] 这是一份衍生合约. 一旦合约签订，就是一份独立的证券，可以单独进行交易. 要进行远期合约的交易，交易者不必拥有真正的黄金. 实际上，这类合约甚至可以建立在只是名义上存在的标的资产上. 关于资产的衍生品亦是如此.

[3] 这样的陈述背后有一些假设，例如期权合约不存在违约风险.

意思，我们要用到极限的概念．令

$$t \to T \tag{4.4}$$

注意在该过程中，

$$\lim_{t \to T} e^{r_t(T-t)} = 1 \tag{4.5}$$

一个问题是 r_t 在极限中出现是否合理．现实中 r_t 和 S_t 都是随机变量，你可能会问常规的极限概念是否仍然适用．现暂时忽略这一顾虑，将该极限带入(4.3)的左端，可得

$$S_T = F(S_T, T) \tag{4.6}$$

由此可知，在到期日，标的资产的价格与远期合约的价格相等．

这是一个边界条件的例子．在到期日——也就是时间变量 t 的边界，定价函数 $F(S_t, t)$ 等于一个特定的值 S_T，该边界条件在 t 时就已经知道，不过 S_t 在 T 的值是未知的．

4.2.2 期权

推导非线性资产的定价函数 $F(S_t, t)$ 并不像远期合约那样简单．在本书后面我们将会作进一步的介绍．这里我们仅介绍一种非线性产品必须满足的重要性质，为接下来的数学工具的学习打好基础．

假定 C_t 是股票 S_t 的看涨期权在 t 时刻的价格．令 r 为无风险利率且为常数，K 是执行价格，T 是到期日，$t<T$. 看涨期权价格可表达为[㊀]

$$C_t = F(S_t, t) \tag{4.7}$$

期权定价函数 $F(S_t, t)$ 有一个基本性质. 在简化条件下，S_t 是影响期权价格的唯一不稳定因素．因此 S_t 不可预测的移动可以引起 C_t 的移动，一旦 S_t 的路径得到确定，就对 $F(S_t, t)$ 的变化施加了一些条件.

为了进一步理解该性质，请看图 4-1. 下半部分为 S_t 的空头方收益图．设一单位的标的资产 S_t 按价格 S 借入或卖出．

图 4-1 的上半部分展示了 S_t 的看涨期权价格函数 $F(S_t, t)$，我们先不考虑 $F(S_t, t)$ 是如何得到和绘图的[㊁].

假定初始时刻标的资产价格为 S，即

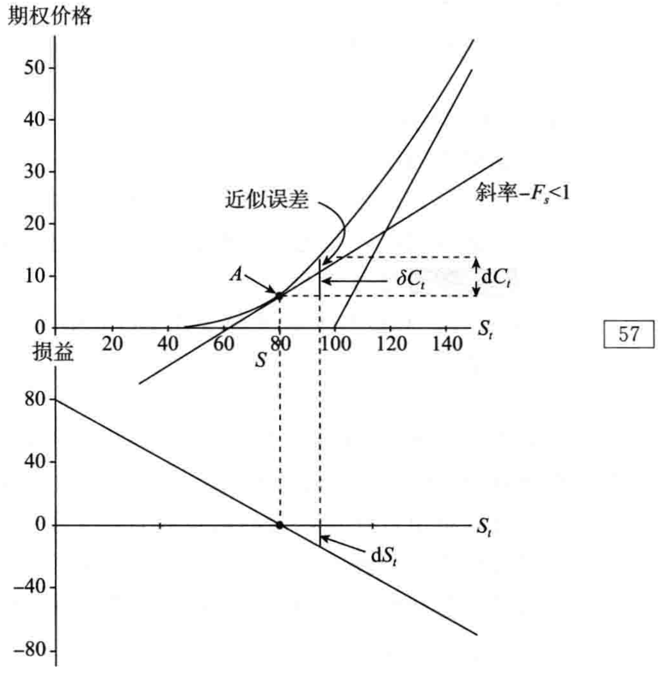


图 4-1 股票价格的变化与看涨期权价格的变化

57

㊀ 利率 r 是常数，所以不影响 $F(\cdot)$.

㊁ 它来自于我们后面介绍的 Black-Scholes 公式．

初始时刻期权价值位于 $F(S_t,t)$上的点 A. 如果股价上升 dS_t，空头方的损失恰为 dS_t，但期权位置上升了.

然而，请注意一个关键点. 根据图 4-1，当 S_t 上升 dS_t，看涨期权价格仅上升 dC_t. 该变化值较小，因为曲线斜率小于 1，即

$$dC_t < dS_t \tag{4.8}$$

因此，如果我们买入一份看涨期权卖出一份股票，当股票上升 dS_t 时，我们会遭受损失.

但这一结果说明通过适当调整持仓量，这种损失可以消除. 考虑 $F(S_t,t)$在 A 点的切线，其斜率为

$$\frac{\partial F(S_t,t)}{\partial S_t} = F_S \tag{4.9}$$

现假设我们卖空 F_S 单位的标的股票. 那么如果 S_t 上涨 dS_t，卖空造成的损失为 $F_s dS_t$. 但根据图 4-1，这一数目与 dC_t 非常接近，可以用 ∂C_t 来近似.

58

由此可见，如果 dS_t 是一个很小的增量，那么用 ∂C_t 来近似替代真实值 dC_t 是一个很好的选择. 结果就是，期权收益会基本上抵消作为空头方的损失. 这样一个投资组合的变化不再是不可预测的.

因此，可以用 $F(S_t,t)$和 S_t 的增量写出等式

$$d[F_S S_t] + d[F(S_t,t)] = g(t)$$

其中 $g(t)$是关于 t 的完全可预测的函数.[⊖]

如果我们学会如何计算这种微分，以上等式就可以来推导 $F(S_t,t)$封闭形式的公式. 如果封闭形式的公式不存在，就要使用数值方法来确定 $F(S_t,t)$的近似轨迹.

下面我们用一个定义来规范表述本节中讨论的一些概念.

定义 12 通过卖出一份期权买入 F_S 份标的资产来抵消 C_t 的风险，这样的操作叫做 Delta 对冲，称这样的投资组合是 Delta 中性的，参数 F_S 记为 Δ.

注意，当 dS_t 较大时，近似

$$\partial C_t \approx dC_t \tag{4.10}$$

会失效，当 S_t 变化较大时，对冲效果也会变差. 这一点可以通过图 4-2 来理解，如果 S_t 的增量为 dS_t，那么对应的 dC_t 会远远超过损失值——$F_S dS_t$.

可以看出，连续时间的假设在资产定价中起着基础性的作用. 事实上，我们可以在以小的时间单位间隔不断地调整卖空的数量来抵消期权价格变化的影响. 之所以能对投资组合做这样的调整，是因为连续时间的假设以及无交易费用的假设. 如之前所说，当增量较大时，这种近似效果会变差.

4.3 应用：另一个定价模型

本书主要介绍衍生品定价的数学内容，而非资产定价本身. 但关于衍生品定价的一般

⊖ 也可以是问题中其他参数的函数.

方法的讨论也是必不可少的，下面我们举例说明如何运用这些数学工具．

我们用上一节的讨论来说明如何使用偏微分方程进行定价．

1）假设一名分析师观察到了衍生品的当前价格 $F(S_t,t)$ 和标的资产的当前价格 S_t. 现给定一个标的资产的价格变化 dS_t，想要计算衍生品的价格变化 $dF(S_t,t)$.

2）我们将利用第 3 章的概念，回忆一下微分是一种可以用来表示函数很微小变化的工具．在这个例子中，我们已知函数 $F(\cdot)$，且 $F(\cdot)$ 依赖于 S_t 和 t，因此根据普通微积分，我们可以得到

$$dF(S_t,t) = F_S dS_t + F_t dt \quad (4.11)$$

其中 F_i 表示偏导数[⊖]

$$F_S = \frac{\partial F}{\partial S}, \quad F_t = \frac{\partial F}{\partial t} \quad (4.12)$$

$dF(S_t,t)$ 表示全变化．

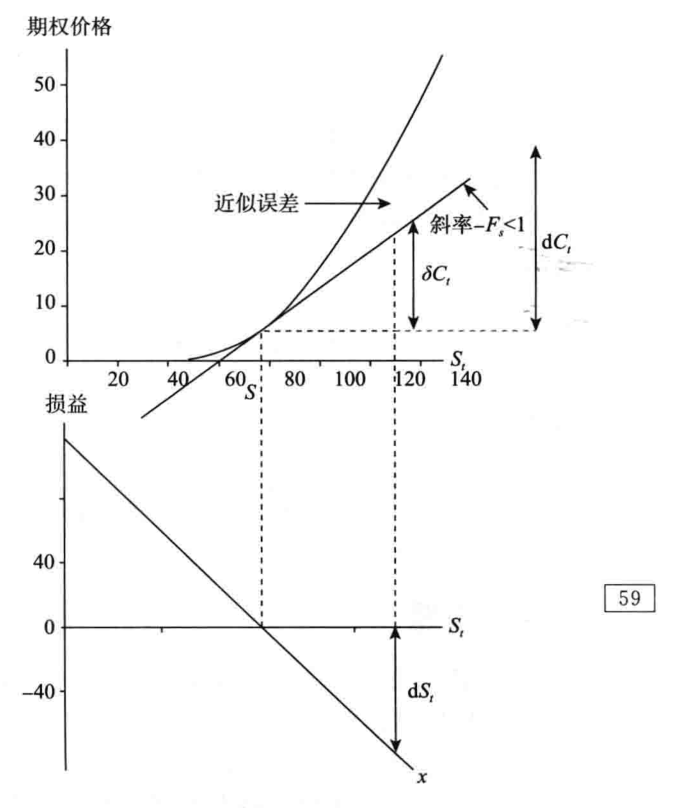


图 4-2　当股票价格移动较大时，股票价格的变化与看涨期权价格的变化

3）等式(4.11)是 $F(\cdot)$ 的全微分，它给出了衍生品价格受其决定因素的影响情况．你可能觉得分析师可以先估计 dS_t，在 F_S、F_t 均有估计值的情况下就可以利用全微分公式(4.11)得到 $dF(S_t,t)$. 但这一切都必须在函数 $F(S_t,t)$ 的形式是已知的且等式(4.11)是正确的情况下才能进行．如果变量都是连续时间的随机过程，还能这样处理吗？回答是否定的，但是，借助随机微积分这个新的工具，我们也将能处理这个问题．

59

4）一旦随机版的(4.11)式已经确定，我们就可以按以下方法完成定价工作．通过 Delta 对冲和无风险投资组合，我们可以得到 $dF(S_t,t)$、dS_t 和 dt 之间的关系式，继而求出类似(4.11)中的微分值．

5）我们可以得到 $F(\cdot)$ 的偏导数之间的关系．这类公式通常为一个偏微分方程．如果 $F(S_t,t)$ 存在解析解，只要有足够的边界条件，我们就可以解出 $dF(S_t,t)$. 在这类问题中未知量是函数，这也说明偏微分方程及它们的解是非常有研究价值的．

60

再举一个例子．

⊖　注意它与 $F(S_t,t)$ 的区别，后者代表衍生品在 t 时刻的价格，而 F_t 表示 $F(S_t,t)$ 关于 t 的偏导数．

例子

假设 $F(x)$关于 $x\in[0,X]$的偏导数是一个已知的常数

$$F_x = b \tag{4.13}$$

该等式是最简单的偏微分方程，它给出了 $F(x)$的偏导数，而 $F(x)$的具体形式未知．

通过该偏微分方程，我们能否写出函数 $F(x)$的具体形式吗？回答是肯定的，因为只有线性函数才可以得到(4.13)．因此，$F(x)$可写为

$$F(x) = a + bx \tag{4.14}$$

$F(x)$的形式已确定．但参数 a 仍未知，可以用"边界条件"来求出．

例如，如果在边界 $x=X$，有

$$F(X) = 10 \tag{4.15}$$

那么 a 可写为

$$a = 10 - bX \tag{4.16}$$

一般在衍生品定价中，必然能找到 $F(\cdot)$在到期日的信息．结合偏微分方程，这些信息可以用来确定 $F(\cdot)$的具体形式．

4.4 问题

前面讨论的解决方法看起来很抽象，但实际上是一个直截了当的方法．然而，还有一个基础性的问题．[⊖]

来自金融市场的数据都是非确定性的．事实上，我们研究的所有关于 t 的变量几乎都是随机的．由于时间连续，随着时间推移，我们会观察到许许多多随机变量的变化．所以 $F(S_t,t)$、S_t 甚至无风险利率 r_t 在连续时间下可能都是随机过程．

我们能否基于同样的理论，用普通的微积分法则得到

$$\mathrm{d}F(t) = F_S\mathrm{d}S_t + F_r\mathrm{d}r_t + F_t\mathrm{d}t \tag{4.18}$$

吗？答案是否定的．当变量是随机过程时，我们需要新的微积分法则和公式，以下是这些难题的初步介绍．

伊藤引理初探

在普通的微积分中，变量都是确定性的．因此，要得到

$$\mathrm{d}F(t) = F_S\mathrm{d}S_t + F_r\mathrm{d}r_t + F_t\mathrm{d}t \tag{4.19}$$

我们只需使用全微分．$F(\cdot)$的变化与其他变量的关系由(4.19)右端给出．但根据相应定义，该等式只有在增量无穷小时才恰好成立．在非无穷小的区间上，等式(4.19)有误差．

⊖ 其实现在有两个问题．还有一个是，已知

$$\mathrm{d}F(t)=F_S\mathrm{d}S_t+F_r\mathrm{d}r_t+F_t\mathrm{d}t \tag{4.17}$$

我们还不知道如何利用套利去掉 $\mathrm{d}t$、$\mathrm{d}F_t$、$\mathrm{d}S_t$、$\mathrm{d}r_t$ 等项．以后再讨论该问题．

回顾一元泰勒展开式，设 $f(x)$是关于 $x\in\mathbf{R}$ 的无穷可微函数，可将 $f(x)$在 x_0 附近泰勒展开：

61

$$f(x)=f(x_0)+f_x(x_0)(x-x_0)+\frac{1}{2}f_{xx}(x_0)(x-x_0)^2+\frac{1}{3!}f_{xxx}(x_0)(x-x_0)^3+\cdots$$

$$=\sum_{i=0}^{\infty}\frac{1}{i!}f^i(x_0)(x-x_0)^i \tag{4.20}$$

其中 $f^i(x_0)$是 $f(x)$关于 x 的 i 阶导数在 x_0 处的值．

我们可以利用以下近似重新表达 $\mathrm{d}f(x)$

$$\mathrm{d}f(x)\approx f(x)-f(x_0) \tag{4.21}$$

$$\mathrm{d}x\approx x-x_0 \tag{4.22}$$

因此，

$$\mathrm{d}F(t)=F_S\mathrm{d}S_t+F_r\mathrm{d}r_t+F_t\mathrm{d}t \tag{4.23}$$

前提是$(\mathrm{d}t)^2$，$(\mathrm{d}S_t)^2$，$(\mathrm{d}r_t)^2$ 以及更高阶项足够小，以至于它们可以在多元泰勒展开中舍去[1]．这样的近似处理使 $\mathrm{d}S_t$，$\mathrm{d}t$，$\mathrm{d}r_t$ 的高阶项不在(4.23)右端出现．

现假设 $\mathrm{d}t$ 是一个很小的确定性的变量，那么我们可以说$(\mathrm{d}t)^2$，$(\mathrm{d}t)^3$ 都是很小的量．但不能对$(\mathrm{d}S_t)^2$ 或$(\mathrm{d}r_t)^2$ 作同样的论断[2]．

首先，$(\mathrm{d}S_t)^2$ 和$(\mathrm{d}r_t)^2$ 在小区间上都是随机的[3]，在 $\mathrm{d}t$ 内方差非零．

这就造成了一个问题，一方面，我们使用连续时间下的随机过程，其方差在 $\mathrm{d}t$ 上非零，所以我们用正数记录$(\mathrm{d}S_t)^2$ 或$(\mathrm{d}r_t)^2$ 的平均值．但在这些条件下，没有理由像在普通微积分中那样继续认为$(\mathrm{d}S_t)^2$ 和$(\mathrm{d}r_t)^2$ 足够小，然后令其等于 0.

所以，在连续随机环境下，我们必须把全微分改写为：

$$\mathrm{d}F(t)=F_S\mathrm{d}S_t+F_r\mathrm{d}r_t+F_t\mathrm{d}t+\frac{1}{2}F_{SS}\mathrm{d}S_t^2+\frac{1}{2}F_{rr}\mathrm{d}r_t^2+F_{Sr}\mathrm{d}S_t\mathrm{d}r_t \tag{4.24}$$

这是我们学习随机微积分的原因之一，我们想知道如何在随机环境下使用链式法则，怎样理解此时的微分．以上例子说明最后的表达式将与普通微积分中的式子有所不同．

如果微分的概念需要改变，那么积分也要重新定义．实际上，在随机环境中我们是用一种新的积分来定义微分，在时间连续的随机环境下，先定义微分是无法实现的．

4.5 小结

现在，一种寻找证券衍生品的“公平市场价格”的方法可以概括为：利用套利，得出一个由未知函数 $F(S_t,t)$的偏导数组成的偏微分方程，然后解该方程，得到含参数的 $F(\cdot)$，利用边界条件，求出参数．

62

本章还介绍了连续时间的金融研究中基础的数学问题．当变量为连续时间下的随机

---

[1] 这将使表达式成为泰勒展开估计式．

[2] 对交叉项$(\mathrm{d}S_t\mathrm{d}r_t)$也可能不正确．

[3] 在无穷小区间，我们会发现这些项的均方极限都是确定值且与 $\mathrm{d}t$ 成比例．

过程时，常规公式将失效，因为这些变量的增量的方差非零，这使得二阶项如$(dS_t)^2$不能忽略．

4.6 参考阅读

Duffie(1996)提供了资产动态估值的详细介绍，Ingersoll(1987)亦是如此．此外还有一些内容较简单的著作，可以帮助读者理解基础的资产定价公式，如 Cox 和 Rubinstein (1985)．大多数定价理论都收录在 Merton(1990)的汇总中．还有部分距今较近一些的著作，更广泛地总结了估值理论，如 Björk(1999)、Nielsen(1999)和 Kwok(1998)．

4.7 习题

1. 假设你关于美国总统选举下注，其中一名候选人是现任总统．市场对你的回报 R 如下：

$$R=\begin{cases}1000\text{ 美元} & \text{若现任总统胜}\\ -1500\text{ 美元} & \text{若现任总统输}\end{cases}$$

你可以任选一边．将现任总统获胜的概率记作 p，$0<p<1$．

(a) 若 $p=0.6$，期望收益是多少？

(b) p 是多少对你如何决定重要吗？

(c) 会有两人按照他们协商一致的 p 估计值往不同方向下注吗？如果有，哪一边是正确的？你能辨别吗？

(d) 统计或经济理论能否帮我们确定 p？

(e) 你会在何种程度上依赖统计学家的话来做你下注的选择？

(f) 你会下多少钱作为赌注？

2. 现在，假设其他条件不变，市场提供的回报 R 如上一题所示．你的一个好朋友向你提供另一个下注方式．回报 R^* 如下：

$$R^*=\begin{cases}1500\text{ 美元} & \text{若现任总统胜}\\ -1000\text{ 美元} & \text{若现任总统输}\end{cases} \tag{4.25}$$

注意 R 和 R^* 对应的是同一个事件．现考虑以下问题：

(a) 利用 R 和 R^*，建立一个下注组合，确保得到无风险回报(假定你的朋友和市场都不会毁约)．

(b) 概率 p 对制定该下注组合重要吗？你关心 p 的值吗？假如你已知 R，而 R^* 在现任总统获胜时的回报尚未确定，你能否根据该下注组合求出未知值的范围？

(c) 统计学家或经济学家在这些决策中起何种作用？为什么？

3. 考虑一种简化的轮盘赌游戏．该游戏只有 2 种结果：红和黑．玩家可以选择一种颜色下注．若选择正确，则每单位赌注可获得 1 美元，若选择错误，则获得 0 美元．假如出现红和黑的概率是一样的，游戏的期望回报是多少？参与者愿意为每次游戏付多少钱？如果允许参与者参与游戏无穷多次，答案会改变吗？

4. 考虑一个公平的抛硬币游戏．如果要出现 3 次正面朝上，抛投的期望次数是多少？写一个模拟程序来验证该结果．

63

第5章　概率论工具

5.1　简介

首先，本章将复习概率论中的一些基本概念，目的是为后面讨论鞅及与鞅有关的知识做准备．我们将介绍随机变量和随机过程的性质．概率论基础较好的读者可以跳过这几节．

其次，本章将介绍二叉树，它在资产定价中起着重要的作用．在连续时间下建立的衍生资产定价模型，往往应用于离散的小区间时间上．例如，以有限差分法、网格法为核心的资产定价方法就属于这一类，它是假定标的资产的价格每隔 Δ 时间观察一次，且在这样的小区间上，[1] 价格只有有限次的变动．这些方法包含一个基本思想：标的资产的价格是连续时间的随机过程，且能够用二叉树模型去近似．本章将介绍这种近似方法及它的合理性．

5.2　概率

衍生品是以标的资产为基础的合约，而标的资产的价格是随机波动的，我们需要用随机数学模型刻画这种随机波动．

一些概率论初等模型特别适用于衍生资产的定价．

让人感到意外的是：目前许多投资者都是凭直觉去理解概率论中的概念，而非公理化的或正规的模型．但是，在第2章的讨论中，我们已经看到，无论"真正的"概率是多少，只要不存在套利机会，我们就能够构造出一个中性概率测度来计算资产的公平价格．所以不管市场参与者的主观想法如何，数学概率模型的效果不会受到影响．

在讨论随机变量前，我们先定义概率空间，这样可以具体地设定框架，使得各种可能性和相应的概率可以正确定义而不至于产生矛盾．

要正式地定义概率模型，我们需要一组基本的未来状态．一个特定的未来状态可以记作 ω．符号 Ω 表示所有可能的未来状态的集合，一次实验的结果取决于是哪一个 ω 发生．从直观上讲，事件是与 ω 相对应的概念．所有事件的集合用符号 I 表示，对每一个事件 $A\in I$，都有一个概率 $P(A)$.

在定义概率时，必须要求满足下列两个条件：

$$P(A) \geqslant 0, \quad 且\ A \in I \tag{5.1}$$

$$\int_{A\in I} \mathrm{dP(A)} = 1 \tag{5.2}$$

第一个条件表明任一事件的概率都是非负的．第二个条件表明所有事件之和的概率为1．注意，这里的记号 $\mathrm{d}P(A)$ 是一个测度符号，可以读作事件 A 的增量概率．

---

[1] 例如，相对于当前值，价格只上升或下降．

三元组$\{\Omega, I, P\}$称为概率空间．其中，ω 是从 Ω 中随机选择的一点，对于 $A\in I$，$P(A)$表示集合 A 包含的所有点所对应的概率之和．

5.2.1 例子

假定交易所交易的某大宗商品期货在某天的价格仅依赖于美国农业部(USDA)在这一天公布的产量报告．

66

USDA 报告的具体数据相当于一个 ω.

报告的数据可能是有利的也可能是不利的，所有有利的数据就构成了一个事件．可能有若干个 ω 对我们开来说都是有利的．从这种意义上来看，一个事件是若干个 ω 构成的集合.

所以如果我们想知道报告有利的概率是多少，可以用以下符号表示：

$$P(\text{产量报告} = \text{有利的}) \tag{5.3}$$

在本例中，Ω 表示 USDA 报告中所有可能会出现的产量数据的集合．

5.2.2 随机变量

一般情况下，概率通常不能用一个简单的数学公式表示出来．不过人们已经建立了一些简洁的数学模型，它们可以用来表示与金融数据相关的事件的概率．[1]

随机变量 X 是一个函数，也是一个 I 到实数集合 $\mathbf{R}$ 的所有子集构成的集族上的映射，对任一个事件 $A\in I$，通过 X 对应 $\mathbf{R}$ 的一个特定子集 B. 因此有

$$X: I \rightarrow B \tag{5.4}$$

在刚才讨论的例子中，生产报告是否有利取决于报告的数据是否达到了一定标准．令 X 为 USDA 将要报告的数值，假定 100 是有利产量的下限，那么可以通过下列关系来理解 X：

$$\text{有利报告} \Rightarrow 100 < X \tag{5.5}$$

其中，假定 X 的取值为实数．

与随机变量 X 及概率相关的一个数学模型是分布函数$G(\cdot)$：

$$G(x) = P(X \leqslant x) \tag{5.6}$$

注：这里 $G(\cdot)$表示 x 的函数．[2]

如果函数 $G(x)$是光滑且可导的，就可以定义 X 的密度函数．密度函数用 $g(x)$表示，具体表达式为

$$g(x) = \frac{\mathrm{d}G(x)}{\mathrm{d}x} \tag{5.7}$$

可以证明在一定的技术条件下，分布函数 $G(x)$总是存在的，但是 $G(x)$能否用公式表

---

[1] 用公式来估计概率是我们接下来将要讨论的重要问题．

[2] 这里 X 代表一个随机变量，而小写的 x 代表某个临界值．

示出来就是另一回事了．在一些著名的模型中，它总是可以表示出来的．下面我们将复习三个在衍生品定价中常用的基本概率模型．

下面的例子都是精心设计的，主要是为了方便读者能理解后面的一些复杂的资产定价方法．首先我们复习一下期望和条件期望的概念．

5.3 矩

有许多方法可以对分布函数进行分类．其中一种用到"矩"的概念．有些随机变量只需前两阶矩就能确定下来，而另一些则需要更多高阶矩才能完全确定．

5.3.1 一阶矩和二阶矩

67

已知随机变量 X 及其密度函数 $f(x)$，其数学期望 $E(X)$叫做一阶矩，定义为

$$E[X]=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x$$

其中 $f(x)$是对应的概率密度函数．[⊖]方差 $E[X-E[X]]^2$ 称为二阶中心矩．随机变量的一阶矩是其分布的"重心"，二阶矩则描述了分布的分散情况．二阶矩的平方根称为标准差，它是衡量观测值与平均值的平均离差的参数．在金融市场中，价格变化的标准差称为波动率.

例如，服从标准正态分布的随机变量 X，其密度函数为

$$f(x)=\frac{1}{\sqrt{2\pi\sigma^2}}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{5.8}$$

其中，方差参数 σ^2 是二阶中心矩，均值参数 μ 是一阶矩．图 5-1 给出了不同参数时正态分布密度函数的图形．

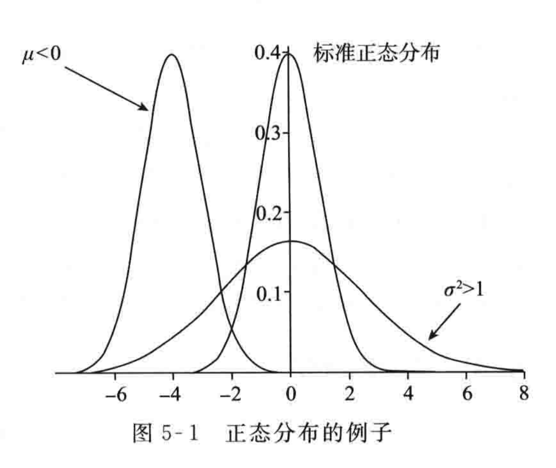


图 5-1 正态分布的例子

随机变量 X 取在任一区间的概率可以通过对密度函数求积分得到．由于 $f(x)$只依赖于两个参数 σ^2 和 μ，因此，若知道了随机变量服从正态分布，则只要有这两个矩的样本估计值，就可以推断出相关的概率值．

正态分布的随机变量 X 也有高阶矩．例如，任一正态分布的随机变量的三阶中心矩满足

$$E[X-E[X]]^3=0$$

68

事实上，所有正态分布的随机变量的高阶矩都可以表示为 σ^2 和 μ 的函数．也就是说，只要正态分布随机变量的一阶矩和二阶矩已知，其余高阶矩不会提供任何额外的信息．

⊖ 如果密度函数不存在，我们用 $\mathrm{d}F(x)$代替 $f(x)\mathrm{d}x$.

5.3.2 高阶矩

考虑图5-2中给出的不对称的密度函数．如果说均值表示重心的位置，标准差刻画分布的离散程度，那么我们还需要一个参数去刻画分布的倾斜程度，三阶矩恰好可以做到这一点．

在金融市场中，厚尾分布是一个重要的概念．图5-2给出了一个对称的密度函数，它与标准正态分布有一个区别，该分布的尾部更厚，这种特性叫"厚尾"，这种分布在金融市场中非常多见．同样，我们需要一个参数来刻画厚尾程度，四阶矩可以做到这一点．

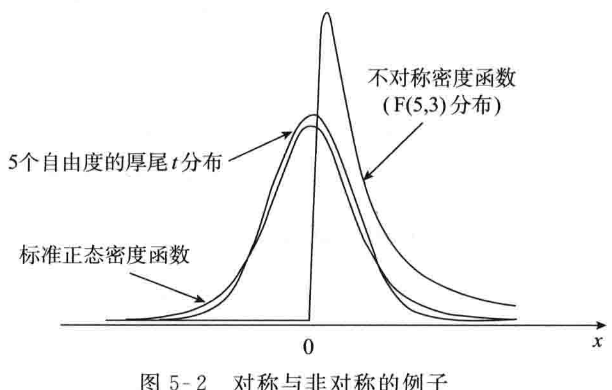


图5-2 对称与非对称的例子

厚尾

厚尾意味着什么?

一个分布尾部比正态曲线更厚，意味着观察到极端值的可能性更大．但作此论断时务必谨慎，正态分布的密度函数的尾部也是延伸到正负无穷大的，所以正态分布随机变量也会随时产生极端值．但相比之下，具有厚尾分布的随机变量极端值可能出现得相对更加频繁一些.

厚尾分布的特性还不止于此．在正态分布中，大多数观测值集中在中心附近．此外，极限值的分布是相对连贯的，从正常值到较大值到极端值，数量逐渐变少．但对于厚尾分布来说，从正常值到极端值出现次数的过渡显得更为快速．对于这样的分布，其尾部中段比正态分布的更瘦，但相比正态分布，它的极端值明显更多．换句话说，一般观察者面对厚尾随机变量的众多极端值，极有可能感到惊讶．

69

5.4 条件期望

所谓"预测"的科学化方法，其实就是测算随机变量的期望值．要预测某个随机变量，我们要利用一些信息，记为I_t. 利用这些信息，算出的期望值叫做条件期望．相应的数学算子称为"条件期望算子".[⊖] 由于不同时间采用的信息通常不同，条件期望算子也有各自的时间下标.

通常，决策者拥有的信息量关于时间递增．如果我们假设决策者不会忘记过去的信息，那么信息集随时间推移而扩张，即

$$I_{t_0} \subseteq I_{t_1} \subseteq \cdots \subseteq I_{t_k} \subseteq I_{t_{k+1}} \subseteq \cdots \tag{5.9}$$

⊖ 算子是一种函数，它将函数映射到函数．也就是说，输入一个函数，将输出另一个函数．

其中 $t_i(i=0, 1, 2, \cdots)$是该信息集可以得到的时刻.

在数学分析中，这样的信息集叫递增的 σ- 域序列. 当这种信息集连续可获得，且满足(5.9)时，信息族 I_t 叫做随指标增加而增加的子 σ- 域族.

条件期望算子能够按以下几个步骤来定义.

5.4.1 条件概率

首先，我们要进一步探讨概率的密度函数. 设 X 是随机变量，$f(x)$为其密度函数，x_0 是 X 的一个可能取值点. 对于充分小的 $\mathrm{d}x$，我们有

$$P\left(|x-x_0|\leqslant\frac{\mathrm{d}x}{2}\right)\approx f(x_0)\mathrm{d}x \tag{5.10}$$

这是 x 落入 x_0 的邻域的概率，邻域的宽度为 $\mathrm{d}x$，如图 5-3 所示.

尽管图 5-3 中的 $f(x)$是非线性的，但在充分小的 $\mathrm{d}x$ 上它近似于直线，图中矩形的面积近似等于 x 落入宽度为 $\mathrm{d}x$ 的邻域的概率. 如果这些概率是建立在信息集 I_t 基础上的，那么 $f(x)$叫做条件密度函数. 由于其依赖于 I_t，所以，记作 $f(x|I_t)$. 如果 $f(x)$不依赖于任何信息集，那么将 I_t 除去，记作 $f(x)$.

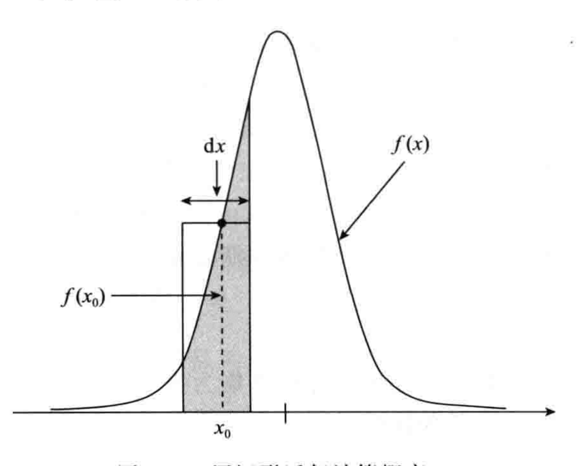


图 5-3 用矩形近似计算概率

看一个简单的例子：股票市场暴跌的几率是一个无条件概率，但如果已知股市已经处于严重的衰退期，那么暴跌的概率就是一个条件概率，衰退已经开始就是已知的信息. 该信息的利用可能导致股市暴跌的条件概率与无条件概率结果大不相同.

条件期望算子

定义条件期望的第二步是考虑“平均”算子. 实际上，每个预测都是根据未来的可能结果得出的一个平均值. 随机变量的估计值是将其未来的可能值根据概率加权平均得到的.

因此，条件期望的计算包含加权求和. 由于可能的结果也许是有限个，也许是无穷多个，当有无穷多个可能结果时，求和要用积分表示.

70

已知时刻 u 的信息集为 I_u，随机变量 S_t 的条件期望可表示为

$$E[S_t|I_u]=\int_{-\infty}^{\infty}S_t f(S_t|I_u)\mathrm{d}S_t,\quad u<t \tag{5.11}$$

在这个表达式中，等式右端可以理解为：对某时刻 t，S_t 所有可能的值根据对应的概率 $[f(S_t|I_t)\mathrm{d}S_t]$ 加权求和. 该平均使用了以 I_u 为条件的条件概率. 如此一来，预测者拥有的信息都被考虑在预测中了.

5.4.2 条件期望的性质

首先介绍一个简便的记号．通常，基于信息集 I_t 得到的条件期望有如下简写形式：

$$E[\cdot \mid I_t] = E_t[\cdot] \tag{5.12}$$

E_t 的下标 t 表明该算子是基于 t 时刻的信息集得到的．

条件期望算子 E_t 有下列性质：

1）两个随机变量和的条件期望等于两个随机变量的条件期望之和：

$$E_u[S_t + F(t)] = E_u[S_t] + E_u[F(t)] \tag{5.13}$$

根据该性质，我们可以分别求随机变量的条件期望再求和，从而得到随机变量和的条件期望.

2）假定最新的信息集是 I_t，但我们想求 $E_{t+T}[S_{t+T+u}]$，$t>0$，$u>0$. 也就是说，我们想要以未来的信息为基础作预测．由于信息集 I_{t+T} 在 t 时刻未知，从而条件期望 $E_{t+T}[S_{t+T+u}]$未知，因此 $E_{t+T}[S_{t+T+u}]$本身也是一个随机变量．条件期望的一个性质是：对 S_{t+T+u}关于未来 $t+T$ 时刻的信息集求条件期望后再对现在的信息集求条件期望，得到的结果等于 S_{t+T+u}关于现在的信息集求条件期望，即

71

$$E_t[E_{t+T}[S_{t+T+u}]] = E_t[S_{t+T+u}] \tag{5.14}$$

根据该性质，递归的条件期望总等于以较小信息集为条件的条件期望：

$$E[E[\cdot \mid I_t] \mid I_0] = E[\cdot \mid I_0] \tag{5.15}$$

其中，I_0 包含于 I_t.

最后，如果条件信息集为空，那么我们得到的是非条件期望 E. 因此，E 也具有条件期望算子的性质．

5.5 一些重要的模型

本节将介绍一些随机变量的重要模型，这些模型不仅在理论上很重要，而且在实际资产定价的应用中也非常有用．本节中，我们讨论的概念也将从随机变量拓展到随机过程．

5.5.1 金融市场中的两点分布

考虑一名通过 Reuters、Telerate 或 Bloomberg 等渠道跟踪衍生资产的价格 $F(t)$的实时交易员．衍生资产的价格 $F(t)$是关于时间连续变化的，假定交易者的注意力有限，每隔 Δ 观察一次价格，Δ 是一个较小的时间间隔．另外，假定在任意时间 t，有两种可能：

1）价格可能按以下规律上升：

$$\Delta F(t) = F(t+\Delta) - F(t) = +a\sqrt{\Delta} \tag{5.16}$$

2)或者，价格可能按以下规律下降：

$$\Delta F(t) = F(t+\Delta) - F(t) = -a\sqrt{\Delta} \tag{5.17}$$

其中 $\Delta F(t)$表示时段 Δ 内观察到的价格的改变量．

假设发生其他结果的概率很小，可以忽略．

那么对于固定的 t 和 Δ，$\Delta F(t)$是一个两点分布的随机变量．$\Delta F(t)$取到各个值的概率分别为

$$P(\Delta F(t)=+a\sqrt{\Delta})=p \tag{5.18}$$

$$P(\Delta F(t)=-a\sqrt{\Delta})=1-p \tag{5.19}$$

时间指数 t 始于 t_0，按 Δ 递增：

$$t=t_0,t_0+\Delta,\cdots,t_0+n\Delta \tag{5.20}$$

每个时间点都会观察到一个 $F(t)$，每个增量 $\Delta F(t)$都取值 $a\sqrt{\Delta}$或$-a\sqrt{\Delta}$. 如果对不同的时间 t，$\Delta F(t)$相互独立，则称增量 $\Delta F(t)$构成的序列为二叉树随机过程，或简称二叉树过程．[㊀]

这些假定是模仿真实市场设定的．但是，在一个高频交易日，可能在许多时间段上 $F(t)$并不改变．也有可能在一些特殊情况下，价格波动大于假定中的值．这些情形我们将在后面再讨论，本章只讨论简单的二叉树．

72

5.5.2 极限性质

在讨论二叉树 $\Delta F(t)$时，很重要的一点是与参数 Δ 有关的两个可能取值．因此，我们要研究二叉树的极限性质，还将讨论一些与衍生品定价相关的问题．

一个重要的问题是：$\Delta F(t)$的典型路径到底是什么样的呢？显然，这样的轨迹是由 $a\sqrt{\Delta}$和$-a\sqrt{\Delta}$组成的序列．如果 $\Delta F(t)$取到每个值的概率都为$\frac{1}{2}$，那么随着 Δ 的减小，$\{\Delta F(t),\ t=t_0,t_0+\Delta,\ \cdots\}$将趋于在 $a\sqrt{\Delta}$和$-a\sqrt{\Delta}$两点间扰动的非常不规则的轨迹．

事实上，随着 Δ 的减小会发生两个结果．第一，观察的时间点会更加靠近；第二，$\left|a\sqrt{\Delta}\right|$会变得更小．

$\Delta F(t)$是价格的增量．那 $F(t)$本身的路径是怎样的呢？首先注意到 $F(t)$表示 t 时刻的价格，那么 $F(t)$等于自 t_0 起到 t 时刻为止原价与所有价格波动之和．当 $\Delta\to 0$ 时，$F(t)$可表示为

$$F(t)=F(t_0)+\int_{t_0}^{t}\mathrm{d}F(s) \tag{5.21}$$

该式的含义是，从初始价格 $F(t_0)$开始，将所有无穷小的变化累加起来，就得到了 t 时刻的价格．显然在连续时间下，价格波动的次数是无穷多的，所以我们要用积分．此外，在极限意义下，用 $\mathrm{d}F(t)$代替了 $\Delta F(t)$，表示变化是无穷小的．

最后，考虑下面的问题：从极限的角度看，无穷小时段上的变化 $\mathrm{d}F(t)$是非常不规则的. $F(t)$的路径会是有界变差的吗?[㊁]这是一个非常重要的问题．如果答案是否定的，那么

㊀ 记住，随机过程是一列以时间为下标的随机变量．

㊁ 参照第 3 章．

黎曼-斯蒂尔切斯积分就不适用于该问题，则需要重新定义一种新的积分．

另一个问题是：(5.21)是关于随机过程的积分，而不是像常规微积分中的那种关于确定性变量的积分，显然，这种积分本身也是随机变量．这样定义积分会有问题吗？我们能不能按照黎曼-斯蒂尔切斯积分的思路构造一种包含随机过程的积分？第9章中的伊藤积分将回答这个问题．

5.5.3 矩

下面我们讨论二叉树的矩．

设 t 为定值，$\Delta F(t)$的数学期望和方差定义如下：

$$E[\Delta F(t)] = p(a\sqrt{\Delta}) + (1-p)(-a\sqrt{\Delta}) \tag{5.22}$$

$$V[\Delta F(t)] = p(a\sqrt{\Delta})^2 + (1-p)(-a\sqrt{\Delta})^2 - E[\Delta F(t)]^2 \tag{5.23}$$

如果在任意时间 t，价格上升和下降的概率均为$\frac{1}{2}$，即

$$p = \frac{1}{2} \tag{5.24}$$

那么 $\Delta F(t)$的期望值为 0，方差为 $a^2\Delta$.

二叉树的方差与 Δ 成正比，这一点很重要．当 Δ 趋于 0 时，方差将按相同速度趋于 0. 这意味着如果我们把 Δ 视为一个很小但不能省略的量，那么方差也是不能省略的．

73

相比之下，如果 $\Delta F(t)$的取值为 $a\Delta$ 或 $-a\Delta$，那么方差将与 Δ^2 成正比．当 Δ 很小时，Δ^2 将会小得更多．当 $\Delta\to 0$ 时，方差将更快地趋于 0，在这种情况下，即使 Δ 不可省略，但方差是可以忽略的．

换句话说，如果随机变量的方差与 Δ^2 成正比，当时间间隔很小时，该变量近似于常数.

图 5-4 描述了与 Δ 成比例的方差和与 Δ^2 成比例的方差的变化趋势．后者在 Δ 很小的时候可以忽略．

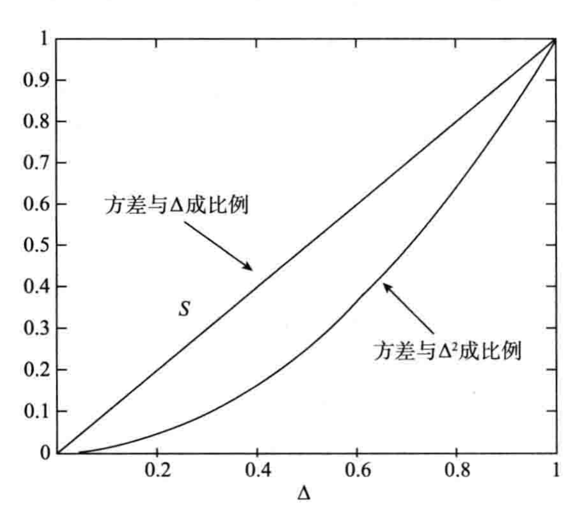


图 5-4 方差与 Δ 成比例与方差与 Δ^2 成比例之间的差异

下面讨论一下高阶矩．同样令 $p=0.5$，由于期望值为 0，那么三阶矩为

$$E[\Delta F(t)]^3 = p(+a\sqrt{\Delta})^3 + (1-p)(-a\sqrt{\Delta})^3 \tag{5.25}$$

再由 $p=0.5$ 知三阶矩为 0.

四阶矩为

$$E[\Delta F(t)]^4 = (+a\sqrt{\Delta})^4 = a^4\Delta^2 \tag{5.26}$$

当 $\Delta\to 0$ 时，四阶矩的值可以忽略，因为它与 Δ 的平方成正比，比时间项 Δ 更快趋于 0.

由以上计算可知，当时间间隔 Δ 很小时，与二项随机变量的高阶矩成正比的量都可以省略．

5.5.4 正态分布

我们继续研究随机变量 $F(t)$，用计算机多次模拟 $F(t)$ 的路径，从同样的初始值开始，得到了多个轨迹．

从 $t_0=0$ 开始，在下一瞬间，$F(t)$ 只有 2 个可能取值：

$$F(0+\Delta)=\begin{cases}F(0)+a\sqrt{\Delta} & \text{概率为 } p\\ F(0)-a\sqrt{\Delta} & \text{概率为 } 1-p\end{cases}\tag{5.27}$$

即 $F(t)$ 在 $t=0+\Delta$ 时刻是两点分布的．

但再经过一段小时间后，比如 $t=2\Delta$，$F(t)$ 就有 3 个可能的取值了，其对应概率如下：

$$F(2\Delta)=\begin{cases}F(0)+a\sqrt{\Delta}+a\sqrt{\Delta} & \text{概率为 } p^2\\ F(0)-a\sqrt{\Delta}+a\sqrt{\Delta} & \text{概率为 } p(1-p)\\ F(0)-a\sqrt{\Delta}-a\sqrt{\Delta} & \text{概率为} (1-p)^2\end{cases}\tag{5.28}$$

也就是说，$F(2\Delta)$ 可能等于 $F(0)+2a\sqrt{\Delta}$，$F(0)-2a\sqrt{\Delta}$ 或 $F(0)$. 如果上升和下降是五五开的话，最后一个取值的概率最大．

现在考虑再过一段时间后 $F(t)$ 的可能取值．不同顺序的上升下降能够产生多种组合．例如，在 $t=5\Delta$ 时刻，一个极端的取值为 [74]

$$F(5\Delta)=F(0)+a\sqrt{\Delta}+a\sqrt{\Delta}+a\sqrt{\Delta}+a\sqrt{\Delta}+a\sqrt{\Delta}\tag{5.29}$$

$$=F(0)+5a\sqrt{\Delta}\tag{5.30}$$

另一个五连降的极端取值为

$$F(5\Delta)=F(0)-a\sqrt{\Delta}-a\sqrt{\Delta}-a\sqrt{\Delta}-a\sqrt{\Delta}-a\sqrt{\Delta}\tag{5.31}$$

大多数取值是既有升也有降的结果．例如

$$F(5\Delta)=F(0)-a\sqrt{\Delta}+a\sqrt{\Delta}-a\sqrt{\Delta}+a\sqrt{\Delta}+a\sqrt{\Delta}\tag{5.32}$$

或

$$F(5\Delta)=F(0)-a\sqrt{\Delta}+a\sqrt{\Delta}+a\sqrt{\Delta}-a\sqrt{\Delta}+a\sqrt{\Delta}\tag{5.33}$$

是两个不同的变化轨迹，它们的最终结果是相等的．

除此之外还有很多可能的取值．事实上，我们能够通过研究一般的情形，计算出 $F(n\Delta)$ 可能取值的总数．显然，当 $n\to\infty$ 时，$F(n\Delta)$ 可能取值的总数也趋于无穷，即使当 $n\to\infty$，$\Delta\to 0$，$n\Delta$ 为常数时，亦是如此，从而我们可以考虑固定时间区间，然后不断地对区间进行越来越细的分割．[⊖] 在 Δ 为常数而 $n\to\infty$ 时，时间是向前推移的，趋于无穷远，

⊖ 其实这种收敛正是我们感兴趣的，这类研究属于弱收敛范畴，它能够告诉我们随机变量在一段固定时间内轨迹的分布情况．

最后的 $F(t)$属于遥远的未来.

问题1：当 $n\to\infty$，Δ 为常数时，随机变量 $F(n\Delta)$服从什么分布？

问题2：当 $n\to\infty$，$n\Delta$ 为常数时，$F(n\Delta)$服从什么分布？[㊀]

回忆一下，$F(t)$一开始是两点分布，随后就变成了多点分布，其概率分布也相应变化. 那么当 $n\to\infty$时，$F(t)$分布的是什么呢？它的极限分布是否存在呢？

这类问题属于“随机变量收敛”问题. 有两种方法可以研究这类问题，一种是中心极限定理；另一种是弱收敛定理.

根据中心极限定理，当 $n\Delta\to\infty$时，$F(n\Delta)$的分布趋于正态分布.

假设 $p=0.5$，且

$$F(0)=0 \tag{5.34}$$

那么当 Δ 固定而 n 较大，$F(n\Delta)$的分布可以用均值为 0、方差为 $a^2n\Delta$ 的正态分布来近似. 它的密度函数为

$$g(F(n\Delta)=x)=\frac{1}{\sqrt{2\pi a^2 n\Delta}}\mathrm{e}^{-\frac{1}{2a^2 n\Delta}(x)^2} \tag{5.35}$$

其对应的分布函数没有解析式，只能用积分表出.

图5-5描述了 $F(n\Delta)$分布的极限形式，结合实际的资产定价来理解其意义可能会更直观. 观察一系列以 n 为参数的随机变量，[㊁]随着 n 的增加，第 n 个随机变量的分布开始接近于正态分布.[㊂]

弱收敛理论将讨论随机变量序列依分布收敛的一系列结论.

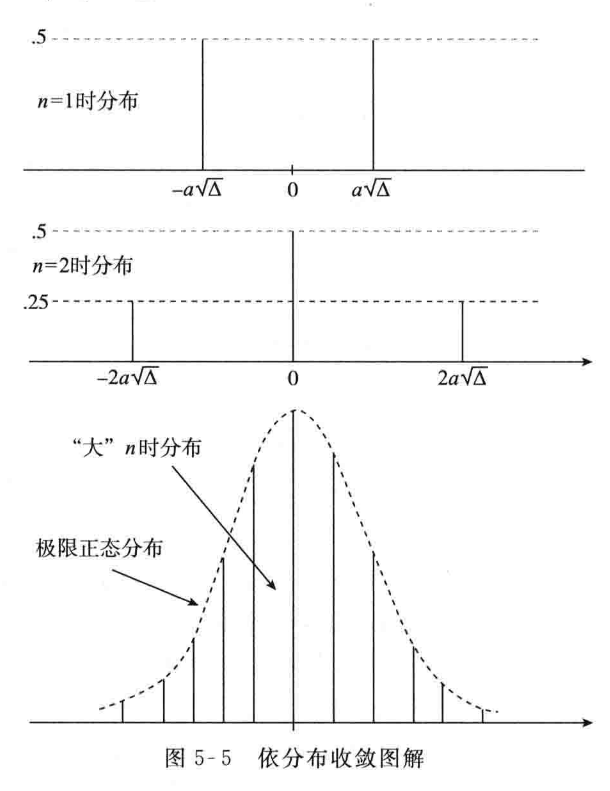


图5-5 依分布收敛图解

75

5.5.5 泊松分布

要研究连续时间的随机过程，我们需要两块奠基石. 一是“连续时间下的正态分布”——布朗运动，也叫维纳过程. 前面已经指出，这些过程的路径一般是连续的.

这意味着，在时间间隔很小时，价格变化也很小的情形，高斯模型是有效的.

㊀ 这里同样，$n\to\infty$.

㊁ 也就是说，我们在观察一个随机过程.

㊂ 再次强调，我们研究的是 $F(n\Delta)$的分布，而不是整个序列$\{F(0), F(\Delta), F(2\Delta), \cdots, F(n\Delta), \cdots\}$.

比如两点分布的极限所讨论的那样情形，随着 $\Delta\to 0$，$\Delta F(t)$也变小，且方差为

$$V(\Delta F(t)) = a^2\Delta \tag{5.36}$$

亦趋于 0. 也就是说，高斯模型要求当时间间隔无穷小时，$F(t)$不能跳跃，变化量$\Delta F(t)$也要为无穷小，且极限为 0.

连续时间下的正态分布(即布朗运动)在资产定价中非常有用. 但有时候，它们不能模拟某些金融市场的价格变化轨迹. 我们还需要一种能模拟价格“跳跃”的模型. 毕竟自从1987 年 10 月股票市场的暴跌之后，有许多价格“跳跃”的实例. [76]

怎样来模拟这种现象呢?

泊松分布是第二块奠基石. 服从泊松分布的随机变量会在不可预知的时刻 $t_i\,(i=1,2,\cdots)$跳跃，这些时刻彼此独立，且跳跃幅度一致. ⊖ 此外在很小的时间区间 Δ，发生两次或以上跳跃的概率忽略不计. t 时间内跳跃的总次数称为泊松计数过程，记为 N_t.

对于一个泊松过程，在小区间 Δ 内发生一次跳跃的概率约为

$$P(\Delta N_t = 1) \approx \lambda\Delta \tag{5.37}$$

其中 λ 为正常数，称为泊松过程的强度.

相比于正态分布，对于一个正态分布的随机变量，它取值为 0 的概率是 0. 而对于泊松分布，若 Δ 很小，该概率约为

$$P(\Delta N_t = 0) \approx 1 - \lambda\Delta \tag{5.38}$$

因此在一个很小的时间段内，很大的概率是没有跳跃发生. 泊松分布的轨迹在大部分区域是连续的，少部分位置由于跳跃而断裂.

一个有限区间 Δ 内发生 n 次跳跃的概率是

$$P(\Delta N_t = n) = \frac{e^{-\lambda\Delta}(\lambda\Delta)^n}{n!} \tag{5.39}$$

它是泊松分布对应的概率分布.

5.6 指数分布

一个参数为 θ 的指数分布的概率密度函数为:

$$f(x) = \frac{1}{\theta}\exp\left(-\frac{x}{\theta}\right), \quad x \geqslant 0 \tag{5.40}$$

积分可得其累积分布函数为

$$F(x) = 1 - \exp(-x/\theta), \quad x \geqslant 0 \tag{5.41}$$

该分布给出了参数为$\frac{1}{\theta}$的泊松分布两次跳跃之间的时间间隔的分布. 易证指数分布的参数 θ 也是它的均值.

图 5-6 给出了两个均值不同的指数分布. 可以看到均值大的指数分布尾部更厚.

⊖ 这些假设都可以调整. 但是，为了维持泊松分布的特征，跳跃的时间必须相互独立.

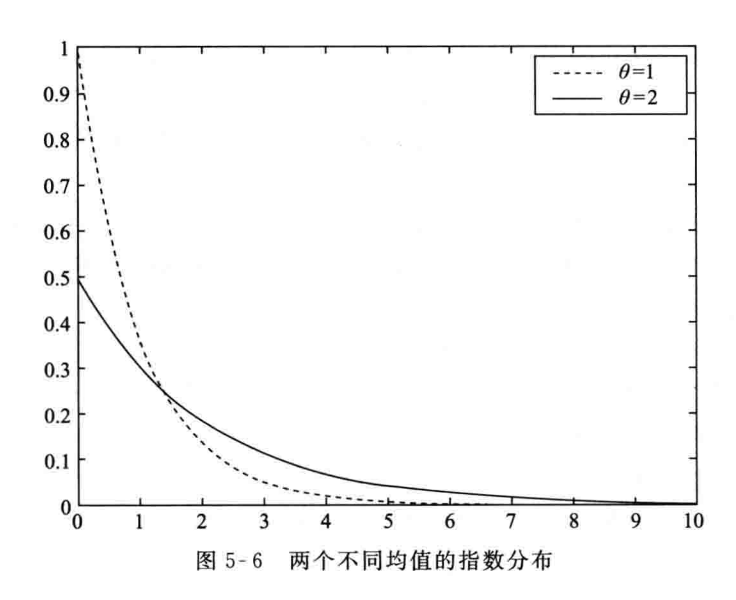


图 5-6 两个不同均值的指数分布

指数分布还有一个有趣的性质，叫做无记忆性．以电话交换台为例，该性质意味着某时间段是否有电话打进来与之前的电话记录无关．

5.7 伽马分布

伽马分布是一个有 2 个参数的分布．它有三种形式的参数模型：

- α 为形状参数，β 为尺度参数；
- α 为形状参数，θ 为比率参数，$\theta=\frac{1}{\beta}$；
- α 为形状参数，μ 为均值参数，$\mu=\alpha\beta$.

以上所有参数均为正实数．伽马分布常用于对等候时间建模．当泊松分布中的事件有相关性时，我们就需要使用泊松分布．

77

服从伽马分布的随机变量具有以下密度函数：

$$f(x)=\frac{1}{\Gamma(\alpha)\beta^{\alpha}}x^{\alpha-1}\mathrm{e}^{-\frac{x}{\beta}} \tag{5.42}$$

其中 α 是形状参数，β 是尺度参数．

该分布与指数分布非常相似，这并不奇怪．因为，如果 α 是正整数，那么 $f(x)$是 α 个服从均值为 β 的指数分布的随机变量之和的密度函数，比率参数 $\theta=\frac{1}{\beta}$. 伽马分布还可以用于建模测量多层泊松回归模型的误差，因为泊松分布与伽马分布可以组合成一个负二项分布. 卡方分布 $\chi^2(d)$是一种特殊的伽马分布，即伽马 $\Gamma\left(\frac{1}{2}d,\ 2\right)$.

随着尺度参数的增大，伽马分布收敛于正态分布．伽马分布的优点是它的取值范围为正实数．

图 5-7 给出了 $\beta=100$，$\alpha=1$ 的伽马分布和 $\mu=10$，$\sigma=10$ 的正态分布的图像. 可以看到两条曲线非常接近. 伽马分布的特征函数为

$$\phi(u)=\left(\frac{1}{1-iu\beta}\right)^{\alpha} \tag{5.43}$$

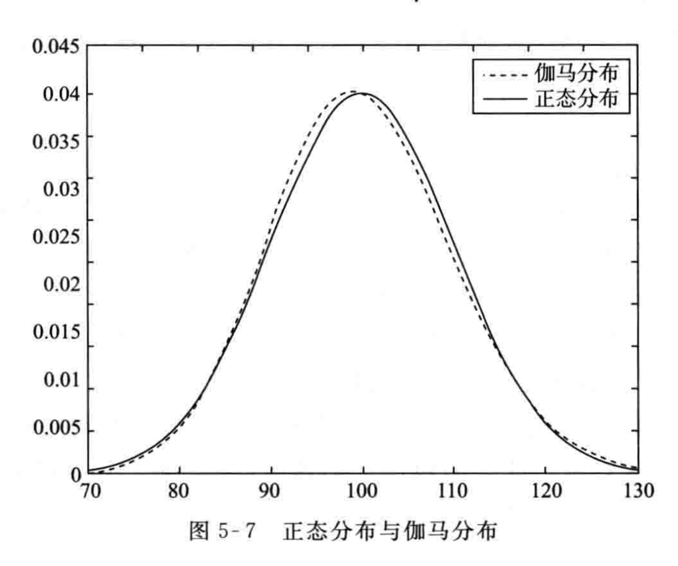


图 5-7 正态分布与伽马分布

5.8 马尔可夫过程及与实际问题的关联

到目前为止，我们已经介绍了随机变量的基础知识 . 但要解决金融市场的问题，仅有这些概念还远远不够，尽管它们是许多复杂模型的基础 . 在金融中，我们需要一种研究随机变量序列的模型，且通常是建立在连续时间下的 . [78]

以 t 为下标的随机变量序列$\{X_t\}$叫做随机过程，t 可以是离散的，$t=0,1,\cdots$，也可以是连续的，$t\in[0,\infty)$. 当 $t\to\infty$时，一般假定随机过程的有限维分布函数

$$F(x_1,\cdots,x_t)=P(X_1\leqslant x_1,\cdots,X_t\leqslant x_t)$$

有定义 . 在 t 为连续时，随机变量可能为不可数个，此时联合分布函数必须谨慎定义，之后将用它来说明维纳过程 .

本节我们将具体讨论一类在衍生品定价中非常有用的随机过程，即马尔可夫过程 . 我们的讨论是建立在离散时间下的，本节将会介绍一些随机过程的重要性质，阐明一些在后面连续时间的利率衍生品模型中会用到的概念 .⊖

定义 13 一个离散时间的随机过程$\{X_1,\cdots,X_t,\ \cdots\}$，其联合概率分布函数为 $F(x_1,\cdots,x_t)$，如果满足

⊖ 金融建模中研究的过程是不是马尔可夫过程非常重要 . 后面章节介绍的 Feynman-Kac 定理只对于马尔可夫过程成立. 但是，可以证明短期利率过程通常不满足马尔可夫条件 . 这就对短期利率过程的数值方法研究增加了限制 .

$$P(X_{t+s} \leqslant x_{t+s} \mid X_t, \cdots, X_1) = P(X_{t+s} \leqslant x_{t+s} \mid X_t) \tag{5.44}$$

那么称该过程为马尔可夫过程，其中 $0<s$，$P(\cdot \mid I_t)$是以信息集 I_t 为条件的条件概率.

马尔可夫过程的假设极具现实意义. 马尔可夫过程$\{X_t\}$是一个离散时间下的随机变量序列. 假设 X_t 是最新观察值，该序列过去的值与 X_{t+s}，$s>0$ 的值完全无关. 也就是说，未来值 X_{t+s}的任何有关信息，只能从 X_t 得到，而不能从更早的值得到. ㊀

5.8.1 关联性

这些概念将怎样帮助市场参与者呢?

假设 r_t 为瞬时的即期利率. 假定 r_t 是马尔可夫的，它意味着未来值 r_{t+s}的可预测信息只与最新一次观察值有关，即条件(5.44)成立. 这样，我们就能进行如下讨论.

我们将利率的变化值分为可预测的和不可预测的两部分.

79

$$r_{t+\Delta} - r_t = E[(r_{t+\Delta} - r_t) \mid I_t] + \sigma(I_t, t)\Delta W_t \tag{5.45}$$

其中 ΔW_t 是方差为 Δ 的不可预测的随机变量，$\sigma(I_t, t)\sqrt{\Delta}$是利率变化量的标准差. 右端第一项表示利率变化量的可预测部分. 第二项表示已知 I_t 时的不可预测部分.

如果 r_t 是马尔可夫过程，且 I_t 只包含 r_t 当前和过去的值，那么以 I_t 为条件的均值和方差是仅与 r_t 有关的函数，可得

$$E[(r_{t+\Delta} - r_t) \mid I_t] = \mu(r_t, t)\Delta \tag{5.46}$$

以及

$$\sigma(I_t, t) = \sigma(r_t, t) \tag{5.47}$$

这几个式子将在第 11 章介绍完随机微分方程之后进一步讨论. 令 $\Delta \to 0$，即可得 r_t 满足的标准随机微分方程:

$$\mathrm{d}r_t = \mu(r_t, t)\mathrm{d}t + \sigma(r_t, t)\mathrm{d}W_t \tag{5.48}$$

有了这样的模型，我们就能继续处理参数 $\mu(r_t, t)$和 $\sigma(r_t, t)$，进而得到可以模拟利率动态变化的模型.

但是，如果利率不满足马尔可夫条件，以上推导便不再适用，因为条件均值和方差可能和过去的信息有关. ㊁

综上所述，马尔可夫假设与衍生品定价关系密切，至少对于利率衍生品是如此.

5.8.2 向量过程

下面讨论多维马尔可夫过程. 继续选取离散时间 t，$t+\Delta$，…以及利率作为研究对象.

我们将看到，虽然两个过程的组合是满足马尔可夫条件的，但其分量，单变量的情形通常不再满足马尔可夫条件.

㊀ 我们不只是在讨论 X_{t+s}的均值和方差的决定因素. 过去的信息不会影响马尔可夫过程的概率行为.

㊁ 而且，如果利率不是马尔可夫过程，偏微分方程和条件期望的重要关联就不能建立. 蒙特卡洛方法与利率衍生品领域的偏微分方程不再等价.

这一点可从固定收益的讨论中看出，我们以收益曲线作为中心话题．所谓的经典方法一般用单一利率过程来模拟利率曲线，如之前讨论的 r_t．另一方面，Heath-Jarrow-Morton (HJM)法是一种更先进的方法，它与 Black-Scholes 理念一致，它通过 k 个独立、联合满足马尔可夫条件的远期利率来建模．但下面我们会看到，k 维马尔可夫过程的单变量过程，不再是马尔可夫过程．所以马尔可夫条件在 HJM 方法下成立，在只研究短期利率时不成立．

假设现有双变量过程，$[r_t, R_t]$，其中 r_t 为短期利率，R_t 为长期利率．它们联合起来是马尔可夫过程：

$$\begin{bmatrix} r_{t+\Delta} \\ R_{t+\Delta} \end{bmatrix} = \begin{bmatrix} \alpha_1 r_t + \beta_1 R_t \\ \alpha_2 r_t + \beta_2 R_t \end{bmatrix} + \begin{bmatrix} \sigma_1 W_{t+\Delta}^1 \\ \sigma_2 W_{t+\Delta}^2 \end{bmatrix} \tag{5.49}$$

其中 $W_{t+\Delta}^1$、$W_{t+\Delta}^2$ 是相互独立的误差项，且它们与过去的 W_s^1、$W_s^2(s \leqslant t)$ 都独立．$\{\alpha_i, \beta_i, \sigma_i\}$ 是常系数．由系统(5.49)可知，当前的短期和长期利率只与最新一次观测到的 r_t 和 R_t 有关．[⊖]

80

虽然这是一个特例但很能说明问题．我们可以从系统(5.49)中提取 r_t 的单变量模型．提取过程用到了递归法，该方法在解差分方程时经常用到．

为得到单变量模型，先看(5.49)的第一个等式：

$$r_{t+\Delta} = \alpha_1 r_t + \beta_1 R_t + \sigma_1 W_{t+\Delta}^1 \tag{5.50}$$

根据(5.49)的第二个等式得到 R_t 的表达式

$$R_t = \alpha_2 r_{t-\Delta} + \beta_2 R_{t-\Delta} + \sigma_2 W_t^2 \tag{5.51}$$

带入(5.50)，得

$$r_{t+\Delta} = a_1 r_t + \beta_1 [\alpha_2 r_{t-\Delta} + \beta_2 R_{t-\Delta} + \sigma_2 W_t^2] + \sigma_1 W_{t+\Delta}^1 \tag{5.52}$$

重组，得

$$r_{t+\Delta} = \alpha_1 r_t + \beta_1 \alpha_2 r_{t-\Delta} + \beta_1 \beta_2 R_{t-\Delta} + [\beta_1 \sigma_2 W_t^2 + \sigma_1 W_{t+\Delta}^1] \tag{5.53}$$

现在右端有一个 $R_{t-\Delta}$，同样可以用下列等式替代：

$$R_{t-\Delta} = \alpha_2 r_{t-2\Delta} + \beta_2 R_{t-2\Delta} + \sigma_2 W_{t-\Delta}^2 \tag{5.54}$$

不断使用该方法，假定随着 k 增加到一定程度，$R_{t-k\Delta}$ 的系数可忽略，那么我们就得到

$$r_{t+\Delta} - r_t = a_0 r_t + a_1 r_{t-\Delta} + a_2 r_{t-2\Delta} + \cdots + [b_0 W_{t+\Delta}^1 + b_1 W_t^2 + \cdots + b_2 W_{t-\Delta}^2] \tag{5.55}$$

显然，这样的 r_t 过程不是马尔可夫过程．此处 $r_{t+\Delta} - r_t$ 的预测必须用到 r_t 的历史值．因此，满足马尔可夫条件的短期利率不能正确地模拟本问题中的短期利率，尽管在本假设中，短期和长期利率的联合是满足马尔可夫条件的．

所以尽管双变量情况下 r_t 是马尔可夫过程，但单变量情况下它并不满足马尔可夫条件．此性质可以反过来用，一个单变量的非马尔可夫过程在增加了维数后，可能转变为马尔可夫过程．这意味着我们可以设定远期利率是马尔可夫的，同时认为将即期利率设

⊖ 这里 W_t^i 不代表维纳过程．它们可以是任何独立同分布且与过去无关的随机变量．

为马尔可夫过程通常是不准确的．这一点在利率敏感的证券建模中非常重要．在利率曲线的研究中，该性质促使我们把重心放在研究 k 维马尔可夫过程而不是非马尔可夫的单变量模型.

5.9 随机变量的收敛性

随机变量的收敛性在资产定价中有很多用途，不论在理论上或是在应用上都是非常重要的. 前面二叉树的例子说明，收敛的概念可以用于用简单模型近似复杂的随机变量．随着$\Delta\to 0$，近似效果越来越好．本节我们将给出一些这类问题的综合处理方法，当然这些讨论只是一些简单的入门级介绍．

5.9.1 收敛的种类及其用途

在金融证券定价中，至少用到三种常用的收敛．

第一种是“均方收敛”. 伊藤积分就是用这种收敛定义的．该积分在构造随机微分方程中很常用．因此均方收敛在随机微分方程的数值计算中起着基础性的作用．

81

定义 14 令 X_0，X_1，…，X_n，…为一组随机变量序列．若满足

$$\lim_{n\to\infty} E[X_n - X]^2 = 0 \tag{5.56}$$

则 X_n 均方收敛于 X.

根据该定义，若随机近似误差 ε_n 按下列定义

$$\varepsilon_n = X_n - X \tag{5.57}$$

则当 n 趋于正无穷时，ε_n 的方差将越来越小．

注意到，对于有限的 n，ε_n 的方差可能很小，但不一定为 0. 这就意味着我们必须对这种误差具体量化，方法之一是用 ε_n 的标准差作为估计值．

1. 均方收敛与其他收敛的关联

均方(m. s.)收敛是一个重要的概念，因为伊藤积分是通过均方极限定义的．如果采用其他收敛来定义，该积分可能不存在．

下面我们进一步讨论这一点，首先考虑一种普通微积分中极限定义的拓展．

定义 15 称随机变量 X_n 几乎必然收敛于 X，如果对于任意 $\delta>0$，有

$$P\left(\left|\lim_{n\to\infty} X_n - X\right| > \delta\right) = 0 \tag{5.58}$$

该定义是普通微积分中极限定义的拓展．它意味着当 n 趋于无穷，两个随机变量的差小到可以忽略．在均方收敛中，是方差收敛到 0，而此处是 X_n 与 X 的差收敛于 0. 在极限处，两个随机变量几乎相同．

2. 例子

设 S_t 是等距时间点上观察到的资产价格：

$$t_0 < t_0 + \Delta < t_0 + 2\Delta < \cdots < t_0 + n\Delta = T \tag{5.59}$$

对每个下标 n，定义随机变量 X_n：

$$X_n = \sum_{i=0}^{n} S_{t_0+i\Delta}[S_{t_0+(i+1)\Delta} - S_{t_{0+i\Delta}}] \tag{5.60}$$

这里$[S_{t_0+(i+1)\Delta} - S_{t_0+i\Delta}]$表示资产价格在 $t_0+i\Delta$ 时刻的增量．从 t_0 开始，每隔 Δ 分钟记录一次．

注意到 X_n 与黎曼-斯蒂尔切斯积分中的求和类似，可看作将区间$[t_0,T]$分成 n 等份，X_n 近似等于

$$\int_{t_0}^{T} S_t \mathrm{d}S_t \tag{5.61}$$

但这里存在一个本质的区别，和式 X_n 中包含了随机过程．所以在作极限得到(5.61)的过程中，必须使用一种新型的收敛方法，常规定义下的极限不再适用．

哪种(随机)收敛定义合适呢?

结论是，当 S_t 是维纳过程时，X_n 不依概率收敛，[⊖]但它是依均方收敛的．所以不同的近似方法是有区别的，关于这部分内容，我们将在伊藤积分的章节中继续讨论．

5.9.2 弱收敛

依均方收敛的概念可以用于寻找随机变量的近似值．当参数 n 趋于正无穷，随机变量 X_n 的值近似等于 X.

82

在弱收敛中(第三种收敛)，并不是针对随机变量 X_n 的值是否收敛，而是针对与序列 X_0，X_1，…，X_n 相关的概率，弱收敛主要用于近似计算随机变量族的分布函数的极限．

定义 16 设 X_n 为随机变量，n 为下标，其概率分布为 P_n. 称 X_n 弱收敛于 X，记为

$$\lim_{n\to\infty} P_n = P \tag{5.62}$$

其中 P 为 X 的概率分布，如果

$$E^{P_n}[f(X_n)] \to E^P[f(X)] \tag{5.63}$$

对任意有界、连续的实值函数 $f(\cdot)$成立，其中 $E^{P_n}[f(X_n)]$表示 $f(X_n)$在概率分布 P_n 下的期望；$E^P[f(X)]$表示 $f(X)$在概率分布 P 下的期望．

根据该定义，随机变量 X_n 弱收敛于 X 只需它们的函数的期望值非常接近．所以 X_n 和 X 的取值不一定很靠近，但是当 $n\to\infty$时，它们在任意区间上的概率必须非常接近．

1. 弱收敛与实际问题的关联

我们通常对那种当 $n\to\infty$时得到的随机变量值很感兴趣．例如，定义伊藤积分前，首先要构造一个结构简单的随机变量，该随机变量依赖于参数 n. 第二步，证明当 $n\to\infty$时，该变量均方收敛到伊藤积分．

所以在定义伊藤积分时，随机变量的值是具有重要意义的，且将用到均方收敛．

其他时候，这些具体值可能不重要，人们可能只关心期望值，如随机过程的某种平

⊖ 如果 S_t 偶尔发生跳跃，该结论依然成立．

均值.

例如，设 $F(S_T,T)$代表某衍生品在到期日 T 的随机价格. 该衍生品的标的资产价格为 S_T. 我们知道如果无套利机会，存在风险中性概率 Q，使得在一定的简化假设下，时间 t 的衍生品价值可表示为

$$F(t) = e^{-r(T-t)} E_t^Q[F(S_T,T)] \tag{5.64}$$

所以，我们要计算 S_T 的函数 $F(\cdot)$的期望值，而不是 S_T 的值. 通过弱收敛的概念，我们能够得到 S_T 的估计值. 有时候这比直接使用随机变量 S_T 更方便. 例如，S_T 可能是连续时间下的随机过程，也可能是定义在小区间上关于某参数 n 的随机序列. 如果使用计算机，这样做就会更方便. 这种思想比用二叉树近似连续正态分布更早出现.

2. 一个例子

考虑时间区间[0,1]，t 代表一个特定时刻($t\in[0,1]$). ㊀ 假设已知 n 个观察值 ε_i($i=1$, 2, …, n)，它们是服从均匀分布 $U[0,1]$且相互独立的随机变量. ㊁

83

下面定义随机变量 $X_i(t)$，$i=1$, …, n

$$X_i(t) = \begin{cases} 1 & \text{若 } \varepsilon_i < t \\ 0 & \text{其他} \end{cases} \tag{5.66}$$

可见 $X_i(t)$取值为 0 或 1，依赖于 t 和 ε_i 的值.

利用 $X_i(t)$($i=1$, …, n)可定义随机变量 $S_n(t)$：

$$S_n(t) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(X_i(t) - t) \tag{5.67}$$

图 5-8 是 $n=7$ 时的示意图. 可以看到 $S_n(t)$分段连续，在 ε_i 处跳跃.

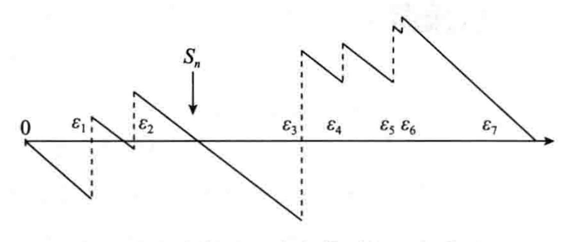


图 5-8 高斯过程在离散时间上的构造

当 $n\to\infty$时，跳跃更加频繁，$S_n(t)$振荡也更加显著. 但跳跃的幅度会减小. 当$n\to\infty$时，对任意 t，$S_n(t)$非常接近于正态分布的随机过程. 有趣的是，该过程的极限是连续的，且起点和终点必为 0. ㊂

当 $n\to\infty$时，$S_n(t)$变得越来越像正态分布过程. 对于较大的 n，我们可以找到一个比 $S_n(t)$更便于使用的极限高斯过程.

需要强调的是，随着 n 的增加，跳跃点的数量也将增加. 在实际应用中，从离散的小区间转变为连续时间来进行分析，这种现象是很常见的.

㊀ 例如，我们可能会令某衍生品合约在时刻 1 到期，而 0 表示现在.

㊁ 这意味着对任意 $0\leqslant t\leqslant 1$,

$$P(\varepsilon_i\leqslant t)=t \tag{5.65}$$

㊂ 这样的过程叫布朗桥.

5.10 小结

本章简要回顾了概率论中一些基础概念.

我们只花很少时间介绍概率的标准定义，但我们介绍了一些别的重要内容.

第一，我们具体介绍了正态分布随机变量和泊松过程.

第二，我们讨论了一个很重要的二叉树. 该例子在许多实际应用中都会出现，且与金融资产定价中很常用的二叉树模型十分相似.

5.11 参考阅读

在本书的余下部分，我们所需要的概率论知识都已包含在本章之中. 但是，对金融市场参与者和金融学习者而言，对随机过程的深入理解将会使其受益匪浅. Ross(1993)是很好的入门读物，Lipster 和 Shiryayev(1977)适合有一定基础的学习者阅读. Cinlar(1978)也适合中等水平的读者阅读. Brzezniak 和 Zastawniak(1999)、Ross(1999)是随机过程的参考文献.

5.12 习题

1. 现有两个离散型随机变量 X, Y，取值为 0 或 1，服从以下联合分布：

	$P(Y=1)$	$P(Y=0)$
$P(Y=1)$	0.2	0.4
$P(Y=0)$	0.15	0.25

84

(a) 求 X 和 Y 的边际分布.

(b) X 和 Y 是否独立?

(c) 计算 $E[X]$与 $E[Y]$.

(d) 计算条件分布 $P[X|Y=1]$.

(e) 求条件期望 $E[X|Y=1]$以及条件方差 $V[X|Y=1]$.

2. 设 X_n 为二叉树，

$$X_n = \sum_{i=1}^{n} B_i$$

其中每个 B_i 相互独立，且具有以下分布

$$B_i = \begin{cases} 1 & \text{概率为 } p \\ 0 & \text{概率为 } 1-p \end{cases}$$

(a) 计算概率 $P[X_4>k]$，$k=0$，1，2，4，画出分布函数.

(b) 计算 X_3 的期望和方差.

3. 设 Z 为指数分布，参数 $\lambda>0$，Z 的分布函数为：

$$P(Z<x)=1-e^{-\lambda x}$$

(a) 确定并画出 Z 的分布函数.

(b) 计算 $E[Z]$.

(c) 求 Z 的方差.

(d) 假设 Z_1 与 Z_2 是相互独立的指数分布，计算它们和的分布：

$$S=Z_1+Z_2$$

(e) 计算 S 的均值和方差.

4. 随机变量 Z 具有泊松分布，只要

$$p(k)=P(Z<k)=\frac{\lambda^k e^{-\lambda}}{k!}, \quad k=0, 1, 2, \cdots$$

(a) 利用展开式

$$e^{\lambda}=1+\lambda+\frac{\lambda^2}{2}+\frac{\lambda^3}{3!}+\cdots$$

证明

$$\sum_{k=0}^{\infty} p(k)=1$$

(b) 计算均值 $E[Z]$和方差 $V[Z]$.

5. 假设随机变量 Z 的定义为 $Z=Z_1+Z_2$，其中 Z_1、Z_2 为独立的泊松分布随机变量，满足

$$p_i(Z_i=k)=P(Z_1=k)=\frac{\lambda_i^k e^{-\lambda i}}{k!} \tag{5.68}$$

其中 $k=0, 1, 2, \cdots, i=1, 2$. 计算 Z 的分布.

6. 利用 $p=0.5$ 的二叉树生成随机游动. 假设随机游动的每步取值为1(概率为 p)或-1(概率为 $1-p$). 生成1000步，计算每一步的期望和方差. 对1000步进行求和. 重复500次，计算均值和方差.

85

第6章　鞅及鞅的表示

6.1　引言

鞅是现代金融理论的核心工具．本章将介绍一些鞅理论的基础知识．鞅理论是一个非常庞大的体系，在此我们只强调与金融衍生品定价直接相关的部分．

我们从符号注释开始讨论．在本章中，我们使用符号 ΔW_t 和 ΔS_t 来表示 W_t 和 S_t 所发生的"小的"变化．有时也使用增量版本 $\mathrm{d}W_t$、$\mathrm{d}S_t$ 来表示在无穷小的时间区间内发生的随机变化．目前，读者可以将这些微分符号理解为在连续时间轴上观测到的"无穷小的"随机变化．这些概念将在第 9 章中进行正式定义．

对于小的时间区间，我们使用符号 h 或 Δ，另一方面，用 $\mathrm{d}t$ 表示无穷小的时间区间．在后面的章节中，我们将证明这些符号并不等价．例如，

$$E[S_{t+\Delta}-S_t]=0$$

的运算是正式定义的，其中 Δ 表示小区间．然而写成，

$$E[\mathrm{d}S_t]=0$$

是不正式的，因为 $\mathrm{d}S_t$ 只是一个符号表示，这一点我们将在伊藤积分的定义中看到．

6.2　定义

鞅理论是根据观测到的发展趋势来对时间序列进行分类．如果一个随机过程的路径没有展示出明显的趋势或周期性，那么它就是鞅．平均而言，呈上升趋势的随机过程被称为下鞅．而上鞅则代表了平均程度上呈递减趋势的随机过程．本节将对这些概念给出严格的定义，首先，我们给出一些符号的定义．

6.2.1　符号

假设我们观察一个以时间 t 为指标的随机变量集族．时间是连续的，我们面对的是连续时间的随机过程，将观察到的过程记作 $\{S_t, t\in[0,\infty]\}$．用 $\{I_t, t\in[0,\infty]\}$ 代表决策者随时间变化可以连续获得的信息集族．[⊖] 当 $s<t<T$ 时，该信息集族满足

$$I_s\subseteq I_t\subseteq I_T \tag{6.1}$$

称 $\{I_t, t\in[0,\infty]\}$ 为滤子．

在讨论鞅理论时(以及本书其他部分)，我们有时候需要考虑随机过程在某些特定时间点上的取值．通常选取一列 $\{t_i\}$，使其满足

⊖ I_t 根据处理的不同问题代表不同的信息．大多数 I_t 的自然使用都是用来表示到 t 时刻为止在金融市场上可获得的价格信息．

$$0 = t_0 < t_1 < \cdots < t_{k-1} < t_k = T \tag{6.2}$$

来表示连续时间区间$[0,T]$内各个时段．注意，这个符号中包含了初值和区间端值，符号t_0表示初始时刻，而t_k是时刻 T 的“新”表示．在这个记号中，当 $k \to \infty$时，$(t_i - t_{i-1}) \to 0$，区间$[0,T]$被分割得更加细小．

现在考虑在有限时间区间$[0,T]$内的随机价格过程 S_t，在特定的时间点 t_i，价格过程的取值为 S_{t_i}．如果对任意 $t \geqslant 0$，S_t 的值都包含在信息集族 I_t 中，那么称$\{S_t, t \in [0,\infty]\}$适应于$\{I_t, t \in [0,T]\}$，即已知信息集 I_t，就可以得出 S_t 的值．

88

现在我们就可以对连续时间的鞅进行定义了．

6.2.2 连续时间鞅

使用不同的信息集可以对价格过程$\{S_t\}$得出不同的预测，这些预测值可以用条件期望来表示．特别地，

$$E_t[S_t] = E[S_T | I_T] \tag{6.3}$$

是在 t 时刻利用已知信息预测得到的S_t的未来值S_T的正式表示．$E_u[S_t]$，$u < t$，表示使用到 u 时刻为止的更小的信息集对相同变量S_t作出的预测．

鞅的性质与这些条件期望相关．

定义 17 如果随机过程$\{S_t,\ t \in [0,\infty]\}$对于$\forall t > 0$ 都满足

1. 当 I_t 已知时，S_t 已知(S_t 关于 I_t 是适应的)．
2. 非条件预测值有限：

$$E|S_t| < \infty \tag{6.4}$$

3. 如果

$$E_t[S_T] = S_t \quad \forall t < T \tag{6.5}$$

的概率为 1，即对于无法被观察到的未来值的最优预测是最近的观察值S_t．

那么称该随机过程为关于于信息集I_t和概率 P 的鞅．

这里所有的期望 $E[\cdot]$，$E_t[\cdot]$都是建立在概率 P 上的．

根据该定义，鞅是在当前信息集的条件下完全无法预测未来变化的随机过程．例如，假设S_t是一个鞅，考虑长度为 $u > 0$ 的时间区间上S_t所发生的变化的预测值：

$$E_t[S_{t+u} - S_t] = E_t[S_{t+u}] - E_t[S_t] \tag{6.6}$$

但 $E_t[S_t]$是对于值已知的随机变量的预测(因为根据定义 $S(t)$是关于I_t适应的)．因此，它等于S_t．如果S_t是鞅，$E_t[S_{t+u}]$也等于S_t．这样就得到了

$$E_t[S_{t+u} - S_t] = 0 \tag{6.7}$$

即对于S_t在任意 $u > 0$ 的时间区间内变化的最优预测值为零．也就是说，鞅未来运动方向是无法预测的，这就是鞅过程的基本特征．如果随机过程的轨迹明显具有可认知的长期或短期趋势，那么该过程就不是鞅．[⊖]

⊖ 一个样本鞅轨迹仍然可以包含看起来像短期存在的趋势．然而，这些向上或者向下的趋势完全是随机的，没有任何系统性特征．

在结束本节之前，我们还要强调鞅定义中非常重要的一点，鞅的定义总是伴随着特定的信息集和特定的概率．如果我们改变信息内容或改变与随机过程相关的概率测度，那么所考虑的随机过程可能不再是鞅．

反之，给定非鞅的随机过程X_t，我们可以通过调整相应的概率测度 P 将X_t转化为鞅．

6.3 鞅在资产定价中的应用

根据上述定义，如果随机过程S_t在给定信息集的条件下完全无法预测未来值，那么它就是鞅．现在我们知道股票价格或债券价格都不是完全无法预测的．贴现债券的价格被认为是随时间递增的，通常股票价格亦是如此．因此，如果B_t表示在 T 时刻($t<T$)到期的贴现债券的价格，那么有$B_t<E[B_u]$，$t<u<T$，显然，贴现债券价格的运动不满足鞅的条件.

89

类似地，一般而言，一支风险股票的价格S_t会一直有着正的预期收益，因此也不会是鞅. 对于小区间 Δ 而言，我们有

$$E[S_{t+\Delta}-S_t]\approx\mu\Delta \tag{6.8}$$

这里 μ 是正的预期收益率．[㊀]

期货和期权也有着相似的结论．例如，期权具有“时间价值”，并且随着时间流逝，假定其他条件不变，欧式期权价格会下降．这种随机过程称为上鞅．[㊁]

如果资产价格更可能是下鞅或上鞅，那么为什么我们还对鞅这么感兴趣呢？

这是因为虽然大多数金融资产的价格不是鞅，但我们可以将它们转化成鞅．例如，我们可以找到一个概率测度 Q 使得债券或股票价格按无风险利率贴现后变为鞅．这种情况下，对于债券以下等式成立

$$E_t^Q[\mathrm{e}^{-ru}B_{t+u}]=B_u,\quad 0<u<T-t \tag{6.9}$$

对于股票价格以下等式成立

$$E_t^Q[\mathrm{e}^{-ru}S_{t+u}]=S_t,\quad 0<u \tag{6.10}$$

这种方法在衍生证券定价中非常有用．

我们将在后面的章节中介绍如何进行这种转化．事实上，有两种方法可以将下鞅转化为鞅．

第一种方法比较直观．我们可以从 $\mathrm{e}^{-rt}S_t$ 或 $\mathrm{e}^{-rt}B_t$ 中减去预期趋势，这会使原有趋势附近的波动完全无法预测，因此，“变形”后的变量是鞅．

这种方法等价于通过分解来得到鞅．事实上，Doob-Meyer 分解意味着，在某些一般条件下，任意连续时间随机过程可以被分解成一个鞅和一个递增(或递减)过程．减去后者即可得到鞅，本章也将介绍 Doob-Meyer 分解．

第二种方法更加复杂，但更加实用．这种方法是改变概率测度，而不是直接减去下鞅.

㊀ 这里的近似在于舍去$E_t[S_{t+\Delta}-S_t]$的泰勒展开式中含有 Δ 的高阶项：$E_t[S_{t+\Delta}-S_t]=\mu\Delta+o(\mu)$，其中 $o(\mu)$表示对应泰勒展开式中所有高阶项．

㊁ 在深度价内的情况下，美式看跌期权可能有负的时间价值．

也就是说，如果有

$$E_t^P[e^{-ru}S_{t+u}] > S_t \quad 0 < u \tag{6.11}$$

这里 $E_t^P[\cdot]$是使用概率测度 P 计算出的条件期望，我们可以尝试找出一个“等价”概率测度 Q，使得在 Q 测度下新的数学期望满足

$$E_t^Q[e^{-ru}S_{t+u}] = S_t \quad 0 < u \tag{6.12}$$

这时$e^{-rt}S_t$就变成了鞅.

将等式(6.11)转化为式(6.12)，新的概率测度称为等价鞅测度，我们将在第 14 章中介绍它.

如果选择第二种方法将任意随机过程转化为鞅，那么我们将用到 Girsanov 定理. 在金融资产定价中，这种方法比 Doob-Meyer 分解有着更加广泛的应用.

6.4 随机建模中鞅的相关知识

当套利机会不存在时，市场均衡意味着我们可以找到一种人造的概率测度 Q 使得所有正确贴现后的资产价格 S_t 变为鞅：

90

$$E_t^Q[e^{-ru}S_{t+u} \mid I_t] = S_t \quad u > 0 \tag{6.13}$$

因此，鞅在资产定价的实践中起着非常重要的作用.

然而这并不是鞅非常有用的唯一原因. 鞅理论内容非常丰富，为讨论连续时间的随机过程提供了良好的环境. 在本节中，我们讨论鞅理论中的一些实用技巧.

用X_t表示在滤子$\{I_t\}$和概率测度 Q 下具有鞅性质的资产价格，它满足

$$E_t^Q[X_{t+\Delta} \mid I_t] = X_t \tag{6.14}$$

其中 $\Delta>0$ 表示小的时间区间. 那么X_t在连续时间下将会有哪种类型的轨迹呢?

为了回答这一问题，首先定义鞅差分 ΔX_t

$$\Delta X_t = X_{t+\Delta} - X_t \tag{6.15}$$

由于X_t是鞅，所以

$$E_t^Q[\Delta X_t \mid I_t] = 0 \tag{6.16}$$

正如前面所说，这个等式意味着无论时间区间 Δ 多么小，鞅的增量都应该是完全无法预测的. 然而，由于我们是在连续时间下讨论的，可以考虑非常小的 Δ. 鞅应该呈现出不规律的轨迹. 事实上，即使时间区间 Δ 无穷小，X_t也不应该呈现出任何可以明显辨认的趋势. 如果它不满足这一点，它就会变成可预测的.

这种不规律的轨迹按两种方式出现. 它们可以是连续的，也可以是跳跃的，前者就是连续鞅，而后者称为右连续鞅.

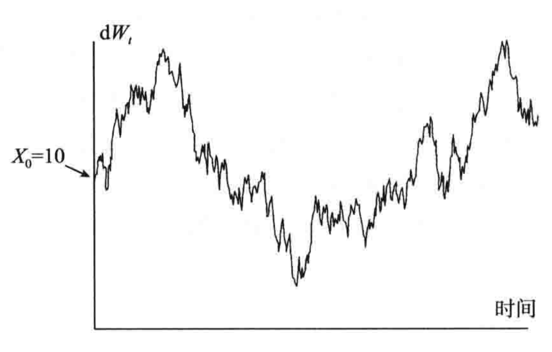


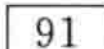
图 6-1 一个连续鞅的例子

图 6-1 给出了一个连续鞅的例子. 注意

它的轨迹是连续的，当 $\Delta\to0$ 时，有

$$P(\Delta X_t > \varepsilon) \to 0, \quad \forall \varepsilon > 0 \tag{6.17}$$

图 6-2 给出了一个右连续鞅的例子. 在这种鞅中，它的路径存在随机的跳跃点，[⊖] 在跳跃时刻 t_0，t_1，t_2，鞅是右连续的(不是左连续的).



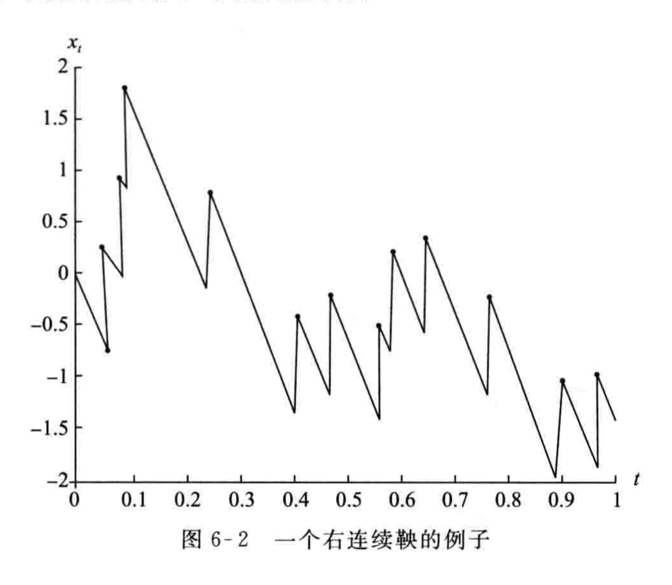


图 6-2　一个右连续鞅的例子

这种路径的不规律性和跳跃的可能性正是表示资产价格时所需要的理论工具，尤其是在已知套利定理的情况下.

除此以外，鞅还有一些重要的含义. 假设我们面对的是连续鞅 X_t，且对任意 $t>0$，X_t 具有有限的二阶矩

$$E[X_t^2] < \infty \tag{6.18}$$

这种随机过程具有有限的方差，被称为**连续平方可积鞅**. 通过调整时间后的布朗运动来对这种鞅进行表示是很有意义的(参见 Karatzas 和 Shreve，1991). 也就是说，连续平方可积鞅非常接近于布朗运动. 这意味着变化的不可预测性和不存在跳跃是连续时间布朗运动的两个重要性质.

这一点本质上意味着，如果连续平方可积鞅适合于对资产价格进行建模，那么可以假设资产价格过程中的微小增量是具有正态性的.

一个例子

在本小节中，我们将要使用两个在“小区间”Δ 内观察到的相互独立的泊松过程来构建一个鞅.

假设金融市场受“好”消息和“坏”消息的影响. 忽略消息的具体内容，仅保留它们是

⊖ 注意这种过程仍然是没有趋势的.

“好”还是“坏”这一信息．N_t^G 和 N_t^B 分别表示到 t 时刻为止出现的“好”消息和“坏”消息的数目．进一步假设消息到达金融市场的方式是与历史数据完全无关的，“好”消息和“坏”消息之间是相互独立的．

92

最后，在小区间 Δ 内，至多出现一条好消息或一条坏消息，并且两种类型消息出现的概率相同．因此，在区间 Δ 内的增量变化 ΔN_t^G 和 ΔN_t^B 的概率分布为

$$P\left(\Delta N_t^G = 1\right) = P\left(\Delta N_t^B = 1\right) \approx \lambda\Delta \tag{6.19}$$

定义变量 M_t

$$M_t = N_t^G - N_t^B \tag{6.20}$$

那么 M_t 是鞅．

为了证明这一点，注意 M_t 在小区间 Δ 内的增量为

$$\Delta M_t = \Delta N_t^G - \Delta N_t^B \tag{6.21}$$

则其条件期望为

$$E[\Delta M_t] = E\left[\Delta N_t^G\right] - E\left[\Delta N_t^B\right] \tag{6.22}$$

由于有近似结果

$$E\left[\Delta N_t^G\right] \approx 1 \times \lambda\Delta + 0 \times (1 - \lambda\Delta) \tag{6.23}$$

$$\approx \lambda\Delta \tag{6.24}$$

$E\left[\Delta N_t^B\right]$同理，因此有

$$E[\Delta M_t] \approx \lambda\Delta - \lambda\Delta = 0 \tag{6.25}$$

因此，M_t 的增量在信息集 I_t 下是无法预测的．可以看出 M_t 满足鞅的其他(数学上的)要求．比如，在 t 时刻，我们知道已经发生的“好”消息或“坏”消息，因此 M_t 是关于 I_t 适应的．

所以只要区间 Δ 内“好”消息和“坏”消息发生的概率(对 ΔN_t^G 和 ΔN_t^B 而言)能够用 $\lambda\Delta$ 的形式表示出来，随机过程 M_t 在 I_t 和这些概率下就是鞅．

然而，如果我们假设“好”消息发生的概率稍大于“坏”消息

$$P\left(\Delta N_t^G = 1\right) \approx \lambda^G\Delta > P\left(\Delta N_t^G = 1\right) \approx \lambda^B\Delta \tag{6.26}$$

那么 M_t 将不会再是 I_t 下的鞅，因为

$$E\left[\Delta N_t^G\right] \approx \lambda^G\Delta - \lambda^B\Delta > 0 \tag{6.27}$$

(事实上 M_t 是下鞅)．因此，改变概率或信息集可能改变随机过程的鞅特征．

6.5 鞅的路径性质

本节我们将进一步介绍连续平方可积鞅的性质．

假设$\{X_t\}$是连续平方可积鞅，选择时间区间$[0,T]$，考虑时间序列$\{t_i\}$：

$$t_0 = 0 < t_1 < t_2 < \cdots < t_{n-1} < t_n = T \tag{6.28}$$

定义轨迹变差：

$$V^1 = \sum_{i=1}^{n} |X_{t_i} - X_{t_{i-1}}| \tag{6.29}$$

V^1可以认为是X_t在时间$[0,T]$内的轨迹长度．

二次变差定义为

$$V^2 = \sum_{i=1}^{n} |X_{t_i} - X_{t_{i-1}}|^2 \tag{6.30}$$

类似地，还可以定义高次变差．例如，四次变差可以被定义为

93

$$V^4 = \sum_{i=1}^{n} |X_{t_i} - X_{t_{i-1}}|^4 \tag{6.31}$$

显然，V^1 或者 V^2 是 X_t 随时间变化的两种不同度量方法．V^1 代表观测到的 X_t 在子区间$t_i - t_{i-1}$内变化绝对值之和，而 V^2 表示变化的平方和．

当 X_t 是连续鞅时，V^1、V^2、V^3、V^4都有一些非常重要的性质．

下面我们回顾一些要点：我们希望X_t是连续的并且方差非零，正如前面提到的，这意味着两点：首先，随着对区间$[0,T]$划分得越来越细，各个“连续的”X_t的值变得越来越接近，对任意 $\varepsilon>0$，如果对所有 i 都有 $t_i \to t_i - 1$，那么

$$P(|X_{t_i} - X_{t_{i-1}}| > \varepsilon) \to 0 \tag{6.32}$$

其次，当划分越来越细致时，我们仍然有

$$P\left(\sum_{i=1}^{n} |X_{t_i} - X_{t_{i-1}}|^2 > 0\right) = 1 \tag{6.33}$$

这个等式是成立的，因为毕竟X_t是方差非零的随机过程．

现在考虑 V^1 和 V^2 的一些性质．

首先，要注意即使 X_t 是一个连续鞅，随着$[t_i, t_{i-1}]$越来越小，X_{t_i}越来越接近$X_{t_{i-1}}$，这也并不意味着V^1就接近于零．读者可能认为这很令人吃惊．毕竟V^1是增量之和：

$$V^1 = \sum_{i=1}^{n} |X_{t_i} - X_{t_{i-1}}| \tag{6.34}$$

当X_{t_i}越来越接近$X_{t_{i-1}}$时，V^1不会也趋近于零吗？

令人吃惊的是，V^1不仅不会趋近于零，还会趋于无穷大．随着$[0,T]$被划分成越来越细的子区间，X_t所发生的变化越来越小．但是，与此同时，定义V^1的区间数也在增加．可以证明，在连续鞅的情况下后者起着主导性的作用，因此V^1会趋于无穷大．除鞅为常数的情况外，连续鞅的轨迹是无界变差的．

这一结论可以由下面具有启发性的式子证明．我们有

$$\sum_{i=1}^{n} |X_{t_i} - X_{t_{i-1}}|^2 < \max_i |X_{t_i} - X_{t_{i-1}}| \times \sum_{i=1}^{n} |X_{t_i} - X_{t_{i-1}}| \tag{6.35}$$

不等号右边的部分通过乘以“最大的”$X_{t_i} - X_{t_{i-1}}$构成．[⊖]这意味着

⊖ 符号$\max_i |X_{t_i} - X_{t_{i-1}}|$表示选择所有观察到$X_{t_i}$的增量中最大的那个．

$$V^2 < \max_i |X_{t_i} - X_{t_{i-1}}| V^1 \tag{6.36}$$

对于所有的 i，当$t_i \to t_{i-1}$时，鞅的连续性意味着“连续的”X_{t_i}相互之间靠得越来越近．在极限情况下，有

$$\max_i |X_{t_i} - X_{t_{i-1}}| \to 0 \tag{6.37}$$

根据等式(6.37)，除非V^1非常大，否则V^2将会在某些概率意义下趋于零．然而这是不可能发生的，因为X_t是方差非零的随机过程，即使将$[0,T]$划分得非常精细也会有$V^2>0$. 也就是说一定会有$V^1 \to \infty$.

现在考虑高次变差的同一性质．例如，考虑 V^4，使用与(6.37)中相同的“技巧”，可以得到

94

$$V^4 < \max_i |X_{t_i} - X_{t_{i-1}}|^2 V^2 \tag{6.38}$$

只要V^2收敛于满足一定条件的随机变量，⊖(6.38)不等号右边的项就会趋于零．原因同上. X_t 是连续鞅，并且随着$[0,T]$的划分越来越细，它的增量越来越小．因此，对于任意 i，当$t_i \to t_{i-1}$时：

$$\max_i |X_{t_i} - X_{t_{i-1}}|^2 \to 0 \tag{6.39}$$

这意味着 V^4 趋于零．对于所有次数高于 2 的变差，这个结论都成立．

对于这些结论的规范证明，读者可以参考 Karatzas 和 Shreve(1991). 这里我们总结出鞅的路径的 3 条性质：

1) 一次变差 V^1 在某些概率意义下会趋于无穷大，连续鞅的行为非常不规律．

2) 二次变差 V^2 收敛于某些良好定义的随机变量．这意味着不论鞅轨迹多么不规律，鞅都是平方可积的，并且在非常小的子区间上增量的平方和收敛．这是可能的，因为一个非常小的数的平方是更小的数．因此，尽管增量之和在某些概率意义下会“非常大”，增量平方和并不会出现这种情况．

3) 所有高次变差都会在某些概率意义下收敛于 0. 我们可以理解为高次变差不会包含像 V^1、V^2 那样很多的信息．

这些性质具有重要的含义．首先，我们可以看出对连续平方可积鞅的积分而言，V^1 并不是非常有用的，但 V^2 非常有用．其次，当标的资产过程是连续鞅时，高次变差可以被忽略．

这些主题将会在我们讨论随机环境下的微分与积分时再次出现．还记得黎曼-斯蒂尔切斯积分定义的读者可能已经发现同样的方法不能应用于连续平方可积鞅意义下的积分. 这是由于黎曼-斯蒂尔切斯积分使用了确定性微积分中 V^1 的等价定义，并且考虑的也是将区间划分得越来越细．但是在随机环境下这种极限并不收敛．

相反，随机微积分不得不使用 V^2. 我们将在后面具体讨论这一问题．

⊖ 不趋于无穷大．

定义 $\| d \| = \max_i |t_i - t_{i-1}|$，那么有

$$[X,X]_T = \lim_{\| d \| \to 0} V^2 \tag{6.40}$$

如果 $X(t)$ 是一个跳跃过程，它的二次变差是随机的，这意味着不同于布朗运动（$[W, W](T)=T$），它是路径依赖的．对于$[0,T]$上两个随机过程X_1和X_2的二次变差，首先定义

$$QC(X) = \sum_{i=0}^{m-1} (X_1(t_{i+1}) - X_1(t_i))(X_2(t_{i+1}) - X_2(t_i)) \tag{6.41}$$

那么就会有

$$[X_1, X_2]_T = \lim_{\| d \| \to 0} QC(X) \tag{6.42}$$

6.6 鞅的例子

在本节中，我们考虑连续时间鞅的一些例子．

6.6.1 例 1：布朗运动

假设 X_t 是一个连续时间且路径连续的随机过程，它的增量服从正态分布．这种随机过程被称为（一般）布朗运动．观察每个 t 处对应的X_t的值，X_t所发生的无穷小的变化记作dX_t．假设X_t的增量变化是独立的．

95

在这种情况下，如果 Δ 是一个小区间，Δ 期间对应的增量 ΔX_t服从均值为$\mu\Delta$、方差为$\sigma^2\Delta$的正态分布

$$\Delta X_t \sim \mathcal{N}(\mu\Delta,\ \sigma^2\Delta) \tag{6.44}$$ ㊀

增量的非相关性可以表示为

$$E[(\Delta X_u - \mu\Delta)(\Delta X_t - \mu\Delta)] = 0, \quad u \neq t \tag{6.45}$$

不考虑 X_t这个随机过程定义的严谨性，我们仅考虑一个简单的问题：X_t是鞅吗？

随机过程 X_t是随时间流逝在无穷小增量上的“累积”，即

$$X_{t+T} = X_0 + \int_0^{t+T} dX_u \tag{6.46}$$

假设上述积分定义是合理的，我们可以计算出相应的数学期望．㊁

考虑关于式(6.41)中给出的概率分布下的数学期望，假设到 t 时刻为止关于 X_t的信息都是已知的：

$$E_t[X_{t+T}] = E_t\left[X_t + \int_t^{t+T} dX_u\right] \tag{6.47}$$

但是在 t 时刻，ΔX_{t+T}的未来值是可以预测的，因为在小区间 Δ 上所有变化的数学期

㊀ ΔX_t 的方差与 Δ 成比例的原因目前尚不清楚．例如，下式是否可能发生？

$$V(\Delta X_t) = \sigma^2\Delta^2 \tag{6.43}$$

问题看上去要复杂得多，这将会是下一章的核心内容．

㊁ 我们还没有定义随机增量的积分．

望等于 $\mu\Delta$，也就是说

$$E_t\left[\int_t^{t+T} \mathrm{d}X_u\right] = \mu T \tag{6.48}$$

因此有

$$E_t[\overline{X}_{t+T}] = X_t + \mu T \tag{6.49}$$

显然$\{X_t\}$在式(6.41)所给概率分布和到目前为止关于X_t的信息下不是鞅．

但是，上面的结果为如何利用$\{X_t\}$进行变换得到鞅提供了思路．考虑新的随机过程

$$Z_t = X_t - \mu t \tag{6.50}$$

可以证明 Z_t 是鞅：

$$E_t[Z_{t+T}] = E_t[X_{t+T} - \mu(t+T)] \tag{6.51}$$

$$= E_t[X_t + (X_{t+T} - X_t)] - \mu(t+T) \tag{6.52}$$

也就是说

$$E_t[Z_{t+T}] = X_t + E[(X_{t+T} - X_t)] - \mu(t+T) \tag{6.53}$$

但是与式(6.46)一样，等号右边的 $E[(X_{t+T} - X_t)]$等于 μT. 这意味着：

$$E_t[Z_{t+T}] = X_t + \mu T - \mu(t+T) \tag{6.54}$$

$$= Z_t \tag{6.55}$$

也就是说 Z_t 是鞅．

因此我们可以通过减去一个确定性函数将 X_t 转化为鞅．同时，注意到这一确定性函数是随时间增加而递增的．该结果在更为一般的情况下依然成立．

6.6.2 例 2：平方过程

考虑在区间 Δ 上增量不相关的随机过程S_t：

$$\Delta S_t \sim \mathcal{N}(0, \sigma^2 \Delta) \tag{6.56}$$

其中起始点为

96

$$S_0 = 0 \tag{6.57}$$

定义新的随机变量

$$Z_t = S_t^2 \tag{6.58}$$

可知 Z_t 是等于 S_t 平方的非负随机变量．那么 Z_t 是鞅吗?

答案是否定的，因为 Z_t 增量的平方是可预测的．考虑 Z_t 增量在"小"区间 Δ 上的数学期望：

$$E_t\left[S_{t+\Delta}^2 - S_t^2\right] = E_t[E[S_t - (S_t - S_{t+\Delta})]^2 - S_t^2] = E_t[S_{t+\Delta} - S_t]^2$$

后一个等式成立，是因为 S_t 的增量是与现在和过去的 S_t 无关的．所以交叉乘积项消失了. 但是这意味着

$$E[\Delta Z_t] = \sigma^2 \Delta \tag{6.59}$$

即 Z_t的增量是可预测的．Z_t 不可能是鞅．

但是，使用例 1 中的方法，我们可以利用均值变换将 Z_t 转化为鞅．事实上，下面的等式很容易证明：

$$E_t\left[Z_{t+T}-\sigma^2(T+t)\right]=Z_t-\sigma^2 t \tag{6.60}$$

通过从 Z_t 中减去 $\sigma^2 t$ 就可以得到鞅．

这个例子再次阐释了一个相同的原理．如果一个随机过程不是鞅，那么减去某个"均值"就可以将它转化为鞅．㊀

这就将我们带回了前面的问题．在金融市场中，我们不能期望观察到的风险证券的市场价格等于它按无风险利率贴现后的期望价格，因为会存在风险溢价．因此，当我们使用无风险利率贴现时，任何风险资产的价格都不可能是鞅．但是前面的讨论表明，证券价格可以被转化为鞅，这种转化将使金融资产的定价变得非常方便．

6.6.3 例 3：指数过程

第三个例子更为复杂，但我们在这里只讨论其中一部分．

再次假设 X_t 的定义与例 1 相同，考虑变换

$$S_t=\mathrm{e}^{\left(\alpha X_t-\frac{\alpha^2}{2}t\right)} \tag{6.61}$$

其中 α 是任意实数．假设 X_t 的均值为零，那么该变换能够得到鞅吗？

答案是肯定的，我们会在后面的章节中进行证明．㊁但是，我们需要注意一些奇怪的地方：X_t 本身也是鞅，为什么还需要减去时间函数 $g(t)$

$$g(t)=\frac{\alpha^2}{2}t \tag{6.62}$$

来确保 S_t 是鞅呢？X_t 的增量是可预测的吗？

这些问题的答案涉及随机环境下衍生品定价的方法，我们将在后面的章节中涉及这一点．

6.6.4 例 4：右连续鞅

再次考虑本章中讨论过的泊松计数过程 N_t，因为它是连续时间过程且跳跃数随时间递增．因此，N_t 不可能是鞅．它有清晰的上升趋势．

然而，用 N_t^* 表示的补偿泊松过程

$$N_t^*=N_t-\lambda t \tag{6.63}$$

97

是鞅．显然，N_t^* 的增量也是不可预测的，它是右连续鞅，方差为有限数，且平方可积．

6.7 最简单的鞅

在复杂的利率衍生品定价中会频繁用到一个简单的鞅，下面我们在离散时间背景下对

㊀ 即通过减去一个记为 $g(t)$ 的时间函数．

㊁ 只要了解伊藤引理．

它进行讨论．

考虑概率 P 下的随机变量 Y_T. 我们将会在 T 时刻得知 Y_T 的值．假设随时间流逝，t，$t+1$，…，$T-1$，T，我们得到的信息集记为 I_t，满足

$$I_t \subseteq I_{t+1} \subseteq \cdots \subseteq I_{T-1} \subseteq I_T \tag{6.64}$$

接下来，考虑概率 P 下 Y_T 在不同时刻的预测值 M_t

$$M_t = E^P[Y_T \mid I_t] \tag{6.65}$$

可以证明预测值序列$\{M_t\}$是鞅．即对于任意 $0<s$：

$$E^P[M_{t+s} \mid I_t] = M_t \tag{6.66}$$

这一结果源自条件期望的递归性质，我们在后面的章节中还将会多次看到这一性质．对于任意随机变量 Z，我们有

$$E^P[E^P[Z \mid I_{t+s}] \mid I_t] = E^P[Z \mid I_t], \quad s > 0 \tag{6.67}$$

即对于未来值的最佳预测就是基于现在的信息进行的预测．将这一性质应用于 $Z=[M_{t+s}]$，可以得到

$$E^P[M_{t+s} \mid I_t] = E^P[E^P[Y_T[I_{t+s}] \mid I_t] \tag{6.68}$$

这一点非常简单．

但 M_{t+s}本身也是一个预测值．在(6.65)等号右侧应用(6.64)性质，有

$$E^P[E^P[Y_T \mid I_{t+s}] \mid I_t] = E^P[Y_T \mid I_t] = M_t \tag{6.69}$$

因此，M_t是鞅．

6.7.1 一个应用

上一节所使用的方法在金融领域中有很多应用，下面我们讨论一种常见的情形．

大多数衍生品在到期日 T 的收益都是随机的．许多衍生品在到期前是没有中间支付的．假设这里讨论的就属于这种情形，并且其到期收益依赖于标的资产的价格S_T，收益函数表示为

$$G^T = f(S_T) \tag{6.70}$$

接下来，考虑 1 美元的投资，连续复利利率为 r_s，这笔投资在 T 时刻到期：

$$B_T = e^{\int_t^T r_s \mathrm{d}s} \tag{6.71}$$

这是 T 时刻所获得的收益总和，且如果 r_s 是随机的，那么获得的收益也将是随机的．这里假设 B_T 已知．

最后，考虑相对价格比率 G_T/B_T. 该比率中含有在 T 时刻才可知的随机变量．当我们获得越来越多关于标的资产价格S_t的信息后，就可以计算出该比率的连续条件期望，直到 G_T/B_T 的值在 T 时刻被确切地知道．将这个利用不同信息集计算出的连续条件期望记为 M_t

$$M_t = E^P\left[\frac{G_T}{B_T}\middle| I_t\right] \tag{6.72}$$

其中 I_t 同往常一样，表示 t 时刻可获得的信息集，P 是合适的概率测度．

98

根据这一结果，连续条件期望构成一个鞅：

$$M_t = E^P[M_{t+s} | I_t], \quad s > 0 \tag{6.73}$$

6.7.2 一个评注

假设 r_t 是随机的，G_T 表示 T 时刻无违约风险的纯折价债券的价格．如果 T 是到期日，那么该债券的面值为

$$G_T = 100 \tag{6.74}$$

而 M_t 是概率测度 P 下收益贴现值的条件期望，根据前面的讨论，它是概率测度 P 下的鞅.

那么现在的问题是 M_t 是否是该贴现债券在 t 时刻的无套利价格呢？换句话说，将无违约风险的纯贴现债券价格记为 $B(t,T)$，并假设 $B(t,T)$是无套利价格，会有如下结果：

$$B(t,T) = M_t \tag{6.75}$$

在本书的第二部分我们将会看到，如果条件期望是在 P 测度下计算的，并且如果该测度是真实世界的测度，那么一般来说M_t将不会等于公平价格 $B(t,T)$.

但是，如果计算 M_t所选择的测度是与无套利"等价"的 Q 测度，那么

$$B(t,T) = M_t \tag{6.76}$$

$$= E^Q\left[\frac{100}{B_T}\middle| I_t\right] \tag{6.77}$$

也就是说M_t正是零息债券的价格．

我们将在后面的章节中介绍如何选择等价鞅测度的，因为鞅是动态资产定价中的关键工具. 此外，我们可以选择不同的概率测度来定义M_t，并且它们(在相应的特定测度下)都是鞅，然而，这些鞅中只有一个等于无套利价格 $B(t,T)$.

6.8 鞅表示

前面的例子说明，可以通过减去一个合适的均值来将一类连续时间过程转化为鞅．

在本节中，我们将这些特殊情形进行规范化处理，并对所谓的 Doob-Meyer 分解进行讨论.

下面我们介绍一个例子，这个例子非常重要，原因(至少)有以下 3 点：

第一个原因是实践方面的，通过对连续时间区间进行划分，我们可以得到一个在金融市场上具有实践意义的证券定价方法．

第二点原因是如果我们从这样的框架开始，理解伊藤积分将会更加容易．

第三，这个例子对概率空间及如何为资产价格的各种路径定义概率测度进行了详细的讨论．

6.8.1 例子

假设交易员在时刻 t_i 观察金融资产的价格 S_t

$$t_0 < t_1 < t_2 < \cdots < t_{k-1} < t_k = T \tag{6.78}$$

如果时刻 t_{i-1} 和 t_i 之间的间隔非常小，并且市场是"流动的"，那么该资产价格在区间 $[t_{i-1},t_i]$ 内至多可能出现一次升或降．我们将表述规范化，即在任意瞬间 t_i，只可能发生两种变化：

99

$$\Delta S_{t_i} = \begin{cases} 1 & \text{概率为 } p \\ -1 & \text{概率为}(1-p) \end{cases} \tag{6.79}$$

假设这些变化是相互独立的．进一步地，如果 $p=1/2$，那么 ΔS_{t_i} 的期望将等于 0，除此以外价格变化的均值都不为 0.

有了这些假定，下面我们从构建基本概率空间开始讨论．

观察 k 个不同时点对应 ΔS_t，[1] 我们下面定义概率 p. $\{p,(1-p)\}$ 表示 S_{t_i} 发生不同变化对应的概率，并且 S_{t_i} 有唯一的（边际）概率分布．我们感兴趣的是价格变化序列的概率．换句话说，我们要讨论不同"路径"的概率．[2] 要做到这一点需要构造出一个概率空间．

我们感兴趣的一个典型目标是价格变化的样本路径或轨迹，首先需要构造一个包含所有可能路径的集合，这一空间被称为样本空间．它的元素是由 $+1$ 和 -1 组成的序列．比如，一个典型的样本路径可以是

$$\{\Delta S_{t_1} = -1,\cdots,\Delta S_{t_k} = +1\} \tag{6.80}$$

由于 k 是有限数，给定起始点 S_{t_0}，我们可以通过将增量相加，轻松地确定资产价格的运动轨迹．可以利用这种方法构造所有可能轨迹的集合，即样本空间．

接下来，定义与这些轨迹相关的概率．当价格变化相互独立（k 是有限数）时，很容易做到这一点．简单地将每一次价格变化的概率相乘，就可以得到该序列的概率．

例如，以 t_0 时刻的 $+1$ 开始交替出现直到 t_k 时刻为止的特殊序列 ΔS^*

$$\Delta S^* = \{\Delta S_{t_1} = +1,\Delta S_{t_1} = -1,\cdots,\Delta S_{t_k} = -1\} \tag{6.81}$$

它对应概率（假设 k 是偶数）为

$$P(\Delta S^*) = p^{k/2}(1-p)^{k/2} \tag{6.82}$$

在前 $k/2$ 段时间内连续下降，然后持续增长到 t_k 时刻的轨迹概率与此相同．

由于 k 是有限数，概率空间中可能轨迹的数目也是有限的，因此我们可以确定每一条轨迹的概率．

值得一提的是这样做的原因，k 的有限数很关键，在轨迹数目有限的条件下，我们才可以逐一地为每个轨迹指定概率．金融市场中衍生品定价经常假设 k 是有限的，并利用这一性质生成概率．

另一个简化了这项工作的假设是连续价格变化的独立性，通过这一假设，整个轨迹的概率可以通过各增量对应概率的简单相乘得到．

我们上面的讨论是针对资产价格变化序列的，而衍生证券一般都是以证券价格为标的价格的. 例如，对于以标准普尔 500 为标的的期权，我们感兴趣的是指数的水平，而不是指数的

[1] 主要 k 是有限数这一假设非常重要．

[2] 比如，交易者可能对于资产价格现在的上升或下降的幅度非常感兴趣．

变化.

给定起始价格S_{t_0}后，我们可以根据后续变化轻松地得到资产价格水平：

$$S_{t_k} = S_{t_0} + \sum_{i=1}^{k}\left(S_{t_i} - S_{t_{i-1}}\right) \tag{6.83}$$

注意，由于S_{t_k}是由ΔS_{t_i}的和构成的，(6.79)也可以用来计算S_{t_k}的概率分布．我们只需简单地将产生同一S_{t_k}的不同轨迹的概率相加即可．[㊀]

100

更准确地说，S_{t_k}的最大可能值为S_0+k. 这一值在所有增量ΔS_{t_i}都为$+1$，$i=1$，…，k，时出现，其概率为

$$P\left(S_{t_k} = S_0 + k\right) = p^k \tag{6.84}$$

类似地，S_{t_k}的最小可能值为S_0-k. 它出现的概率为

$$P\left(S_{t_k} = S_0 - k\right) = (1-p)^k \tag{6.85}$$

在这些极端情况下，只有一条轨迹能够使$S_{t_k}=S_0-k$或$S_{t_k}=S_0-k$.

一般来说，资产价格会在这两种极端值之间．在观察的k个增量中，如有m个为$+1$，$k-m$个为$-1(m\leqslant k)$，S_{t_k}的取值为

$$S_{t_k} = S_0 + m - (k-m) \tag{6.86}$$

注意有一些轨迹最终得到的S_{t_k}值是相等的，将这些轨迹对应的概率相加，就可以得到

$$P\left(S_{t_k} = S_0 + m - (k-m)\right) = \mathrm{C}_k^{m-k} p^m (1-p)^{k-m} \tag{6.87}$$

其中$\mathrm{C}_k^{m-k} = \dfrac{k!}{m!\,(k-m)!}$.

这个概率是二项分布的概率，当$k\to\infty$时，该分布收敛于正态分布．[㊁]

S_{t_k}是鞅吗?

(6.83)式包含了过去的每一步价格变化，这样定义的$\{S_{t_k}\}$是鞅吗?

考虑在(6.79)中给出的概率下的数学期望

$$E\left[S_{t_k}\,\middle|\,S_{t_0},\Delta S_{t_1},\cdots,\Delta S_{t_{k-1}}\right] = S_{t_{k-1}} + [(+1)p + (-1)(1-p)] \tag{6.88}$$

其中等号右边第二项是在信息$I_{t_{k-1}}$下未知的增量ΔS_{t_k}的数学期望．显然，当$p=1/2$时，这一项为0，将会得到

$$E\left[S_{t_k}\,\middle|\,S_{t_0},\Delta S_{t_1},\cdots,\Delta S_{t_{k-1}}\right] = S_{t_{k-1}} \tag{6.89}$$

也就是说在由“过去”价格变化所生成的信息集下和特定概率分布下$\{S_{t_k}\}$是鞅．

如果$p\neq 1/2$，$\{S_{t_k}\}$在$\{I_{t_k}\}$下不再是鞅．但是，下面定义的中心过程$\{Z_{t_k}\}$

㊀ 只要标的事件是互不相容的，就可以将概率相加．在这种特殊情况下，根据定义，不同轨迹满足这一点．

㊁ 这是弱收敛的一个例子．

$$Z_{t_k}=\left[S_{t_0}+(1-2p)\right]+\sum_{i=1}^{k}\left[\Delta S_{t_i}+(1-2p)\right] \tag{6.90}$$

$$Z_{t_k}=S_{t_k}+(1-2p)(k+1) \tag{6.91}$$

在I_{t_k}下还是鞅.㊀

6.8.2 Doob-Meyer 分解

考虑在任意 t_i时刻资产价格升的概率略大于降的概率的情形，可以预测观察到的轨迹总体具有上升趋势：

$$1>p>1/2 \tag{6.92}$$

那么，如前所述

101

$$E^P\left[S_{t_k}\middle|S_{t_0},S_{t_1},\cdots,S_{t_{k-1}}\right]=S_{t_{k-1}}-(1-2p) \tag{6.93}$$

即

$$E^P\left[S_{t_k}\middle|S_{t_0},S_{t_1},\cdots,S_{t_{k-1}}\right]>S_{t_{k-1}} \tag{6.94}$$

根据(6.92)，$2p>1$. 这意味着$\{S_{t_k}\}$是下鞅.

如前所述，我们现在有

$$S_{t_k}=-(1-2p)(k+1)+Z_{t_k} \tag{6.95}$$

其中 Z_{t_k} 是鞅. 因此，我们可以将下鞅分解为两部分. 等号右边第一项是递增的确定性变量，第二项是在t_0时刻值为$S_{t_0}+1-2p$ 的鞅.(6.95)的写法是 Doob-Meyer 分解的一种简化情形.㊁

1. 一般情形

对于在连续时间区间上含有限个观察点的过程来说，可以将具有上升趋势的下鞅分解成一个确定性的趋势部分和一个鞅部分. 那么当涉及整个连续时间过程时，可以进行类似的分解吗?

Doob-Meyer 定理回答了这一问题. 我们不加证明地叙述这个定理.

定理 3 如果X_t是集族$\{I_t\}$下的右连续下鞅，$0\leqslant t\leqslant\infty$，且对于任意 t，$E[X_t]<\infty$，那么X_t有如下分解

$$X_t=M_t+A_t \tag{6.96}$$

其中 M_t 是概率 P 下的右连续鞅，A_t 是 I_t 下的可测增过程.

这个定理说明，即使连续观察的资产价格偶尔有跳跃，并且同时具有上升趋势，我们还是可以通过减去一个 t 时刻观察到的过程将它们转化为鞅.

如果原始的连续时间过程不存在任何跳跃，即是连续的，那么得到的鞅也是连续的.

㊀ 可以通过$\{Z_{t_k}\}$的数学期望进行检验，过去的$\{Z_{t_k}\}$的条件期望等于$\{Z_{t_{k-1}}\}$.

㊁ 该项用于连续时间鞅中. 这里讨论的是连续时间区间内离散的划分.

2. Doob-Meyer 分解的应用

将非鞅过程转化为鞅在金融资产定价中非常有用．我们考虑一个简单的例子．

再次假设时间 $t\in[0,T]$是连续的．以资产价格S_t为标的的看涨期权 C_T 在到期日 T 的价值可以用以下函数表示

$$C_T = \max[S_T - K,0] \tag{6.97}$$

也就是说，如果标的资产价格高于执行价格 K，该期权的价值等于该溢价．如果标的资产价格低于 K，该期权的价值为 0.

在更早的时刻 t，$t<T$，C_T的准确值是未知的．但我们可以利用 t 时刻已知的信息I_t预测

$$E^P[C_T|I_t] = E^P[\max[S_T - K,0]|I_t] \tag{6.98}$$

其中数学期望是根据主导价格运动的分布函数求出的．

给出这个预测值后，我们可能会问，$E^P[\max[S_T-X, 0]|I_t]$以无风险利率 r(为常数)贴现后的值

$$C_t = e^{-r(T-t)}E^P[\max[S_T - K,0]|I_t] \tag{6.99}$$

能够给出看涨期权的公平市场价值C_t吗？

答案取决于在I_t、P 下$e^{-rt}C_t$是否是鞅．如果是，那么我们有

$$E^P\left[e^{-rT}C_t\middle|I_t\right] = e^{-rt}C_t \tag{6.100}$$

102

方程两边同时乘以e^{rt}，

$$E^P\left[e^{-r(T-t)}C_T\middle|I_t\right] = C_t \tag{6.101}$$

即$e^{-rt}C_t$是鞅．

但是$e^{-rt}S_t$在真实概率 P 下是不是鞅呢？

正如第 2 章中讨论的那样，假设投资者是风险厌恶的，那么对于一种典型的风险证券有

$$E^P\left[e^{-r(T-t)}\middle|I_t\right] > S_t \tag{6.102}$$

即

$$e^{-rt}S_t \tag{6.103}$$

是下鞅．

但是，根据 Doob-Meyer 分解，我们可以分解

$$e^{-rt}S_t \tag{6.104}$$

得到

$$e^{-rt}S_t = A_t + Z_t \tag{6.105}$$

其中A_t是增过程，且是I_t可测的随机变量，Z_t是I_t下的鞅．

如果函数A_t可以直接得到，我们可以用(6.102)中的分解和(6.99)来得到看涨期权在 t 时刻的公平市场价值．

然而，这种资产定价方法在实践中很少应用．通过改变相应概率测度 P 将资产价格转化为鞅更为方便和简单，而不是减去它们的漂移项．

6.9 随机积分的第一个例子

我们可以利用目前已有的结果来定义一个新的鞅．

$H_{t_{i-1}}$ 表示适应于$I_{t_{i-1}}$的任意随机变量．[⊖] Z_t表示I_t下和特定概率测度 P 下的鞅．定义随机过程

$$M_{t_k}=M_{t_0}+\sum_{i=1}^{k}H_{t_{i-1}}\left[Z_{t_i}-Z_{t_{i-1}}\right] \tag{6.106}$$

那么该随机过程也是I_t下的鞅．

这种表示背后的思想并不难描述．Z_t是鞅，并且具有不可预测的增量．$H_{t_{i-1}}$适应于$I_{t_{i-1}}$，意味着给定$I_{t_{i-1}}$，$H_{t_{i-1}}$为“常数”．那么Z_{t_i}的增量与$H_{t_{i-1}}$也是不相关的．使用这些观察值我们就可以计算出

$$E_{t_0}\left[M_{t_k}\right]=M_{t_0}+E_{t_0}\times\left[\sum_{i=1}^{k}E_{t_{i-1}}\left[H_{t_{i-1}}\left(Z_{t_i}-Z_{t_{i-1}}\right)\right]\right] \tag{6.107}$$

但是Z_{t_i}的增量在t_{i-1}时刻是不可预测的．[⊜] 但它是I_t适应的．这意味着我们可以将“里面的”算子去掉，得到

$$H_{t_{i-1}}\left(Z_{t_i}-Z_{t_{i-1}}\right)=0$$

从而得到

$$E_{t_0}\left[M_{t_k}\right]=M_{t_k} \tag{6.108}$$

因此 M_t 具有鞅性质．

这样定义的M_t是随机积分的第一个例子．问题在于，当 $\sup_i[t_i-t_{i-1}]$趋于 0 时是否能够得到相同的结果．使用类似的方法，我们可以得到如下

$$M_t=M_0+\int_0^t H_u\mathrm{d}Z_u \tag{6.109}$$

103 表示吗？其中 $\mathrm{d}Z_u$表示在 t 时刻给定信息下均值为 0 的无穷小随机增量．

我们将在接下来的几章讨论这种积分是否有意义，还将回答黎曼-斯蒂尔切斯近似体系是否可以用来定义(6.106)中的随机积分．

金融中的应用：交易收益

随机积分在金融理论中有很多有意思的应用，我们将在本节中讨论其中的一种．

⊖ 读者注意，这意味着给定$I_{t_{i-1}}$中的信息，$H_{t_{i-1}}$的值恰好是已知的．

⊜ 记住，$E_0\left[E_{t_{i-1}}[\cdot]\right]=E_{t_0}[\cdot]$.

考虑在交易时间t_i同时对无风险证券和风险证券进行投资的决策者，其中$0=t_0<t_1<\cdots<t_n=T$.

用$\alpha_{t_{i-1}}$和$\beta_{t_{i-1}}$分别表示投资者在交易时间t_i前持有的无风险证券和风险证券数量．显然这些随机变量是关于$I_{t_{i-1}}$适应的．[1] α_{t_0}和β_{t_0}是非随机的初始持有量．用B_{t_i}和S_{t_i}分别表示时刻t_i无风险证券和风险证券的价格．

假设我们考虑的是自融资交易策略，这种策略是在t_i时刻的投资只能使用t_{i-1}时刻的收益．也就是说，它们满足

$$\alpha_{t_{i-1}} B_{t_i} + \beta_{t_{i-1}} S_{t_i} = \alpha_{t_i} B_{t_i} + \beta_{t_i} S_{t_i} \tag{6.110}$$

其中$i=1, 2, \cdots, n$.

根据这种策略，投资者只能在t_i时刻卖出与等号左边数量相等的头寸，然后用所获得的收益来购买α_{t_i}和β_{t_i}单位无风险证券和风险证券．在这种意义下，他今天所做的投资完全通过他过去的投资进行融资．

分别用t_{i-1}，t_{i-2}，…时刻(6.110)的左边部分，利用关系式

$$B_{t_i} = B_{t_{i-1}} + \left[B_{t_i} - B_{t_{i-1}}\right],\ S_{t_i} = S_{t_{i-1}} + [S_{t_i} - S_{t_{i-1}}]$$

进行递归替换，可以得到

$$\alpha_{t_0} B_{t_0} + \beta_{t_0} S_{t_0} + \sum_{j=1}^{i-1}\left[\alpha_{t_j}\left[B_{t_{j+1}} - B_{t_j}\right] + \beta_{t_j}\left[S_{t_{j+1}} - S_{t_j}\right]\right] = \alpha_{t_i} B_{t_i} + \beta_{t_i} S_{t_i} \tag{6.111}$$

其中右边部分是决策者在t_i时刻的交易后的财富值．

对(6.111)进行仔细考察，可以发现左边部分刚好具有与前一节中讨论的随机积分完全相同的结构．这里α_{t_j}和β_{t_j}是$I_{t_{j-1}}$-可测的．

因此，使用随机积分进行建模恰好符合这种投资情况．

6.10 鞅方法与定价

Doob-Meyer 分解是一种鞅表示理论．这种类型的结果初看上去可能非常普通，它可以将任意一个下鞅C_t分解成两部分．一部分是在t时刻的信息下“已知的”趋势，另一部分是相同的信息集和概率P下的鞅．进行了一些技术调整后，这个结论等价于

$$C_T = C_t + \int_t^T D_s \mathrm{d}s + \int_t^T g(C_s)\mathrm{d}M_s \tag{6.112}$$

104

其中D_s在给定信息集I_s后已知的量，$g(\cdot)$是C_s的非预报函数，M_s是信息集$\{I_s\}$和概率P下的鞅．[2]

在本节中，我们将说明该定理是市场上一些常见行为的抽象表示，它给出了金融建模中常用的鞅方法．

---

[1] 在t_i时刻投资者了解他的无风险证券和风险证券头寸情况．

[2] 我们后面将会看到，函数$g(\cdot)$不可预测的本质意味着$g(C_s)$与$\mathrm{d}M_s$是不相关的．

首先，我们介绍选择该方法的原因．

假设我们要为某衍生证券定价，其价格记为C_t．到期日，它的收益为C_T．在第 2 章中我们已经介绍过，适当规范化后的C_t可以结合鞅测度 Q 得到价格公式

$$\frac{C_t}{B_t} = E_t^Q\left[\frac{C_T}{B_T}\right] \tag{6.113}$$

可以证明该等式能够由(6.112)得出．(6.113)可以看作是利用B_t将C_t规范化后，对(6.112)两边同时取条件期望E_t^Q，其中

$$E_t^Q\left[\int_t^T \widetilde{D}_s \mathrm{d}s\right] = 0 \tag{6.114}$$

$$E_t^Q\left[\int_t^T g\left(\frac{C_s}{B_s}\right)\mathrm{d}s\right] = 0 \tag{6.115}$$

这里 $\widetilde{D}$ 是C_t规范化后的趋势，即比率C_t/B_t.

这是得到(6.113)价格公式的一种方法．已知衍生证券价格C_t，如果我们能够写出它的鞅表示，就可以试着找出中性风险下满足(6.114)和(6.115)的规范化处理．我们可以把这个过程作为衍生证券定价的一般方法．

在下一节中我们使用的就是这种方法，我们将首先说明如何利用衍生证券价格C_t得到鞅表示，然后考察这种鞅表示的含义，并对自融资组合的概念进行解释．

6.11 定价方法

在离散时间中，我们用 $h>0$ 表示小的时间区间，并且像前面几节中那样将区间$[t,T]$分成 n 个这种小区间．C_t和S_t分别表示衍生证券和标的资产的当前的价格，其中C_t未知．到期日为 T. 到期时，衍生品的市场价格等于它的收益

$$C_T = G(S_T) \tag{6.116}$$

其中函数 $G(\cdot)$已知，S_T是 T 时刻标的证券的价格(未知).

与(6.112)等价的离散鞅表示由下面的等式给出

$$C_T = C_t + \sum_{i=1}^{n} D_{t_i}\Delta + \sum_{i=1}^{n} g(C_{t_i})\Delta M_{t_i} \tag{6.117}$$

其中 ΔM_{t_i}为 $\Delta M_{t_i} = M_{t_{i+1}} - M_{t_i}$，且 n 满足

$$t_0 = t < \cdots < t_n = T \tag{6.118}$$

鞅表示在确定衍生证券的无套利价格C_t时是如何起作用的呢?

6.11.1 套期保值

105

第一步是为证券价格C_t构造一个“套期保值”.

我们可以利用第 2 章中使用的标准方法做到这一点．用B_t表示短期利率 r 下的无风险借贷(假设 r 为常数)，用S_{t_i}表示标的证券在t_i时刻的价格，因此$[B_{t_i}, S_{t_i}]$在t_i时刻已知．

现在，假设我们选择前几节中的α_{t_i}和β_{t_i}来复制资产组合

$$C_{t_i} = \alpha_{t_i} B_{t_i} + \beta_{t_i} S_{t_i} \tag{6.119}$$

其中 α_{t_i} 和β_{t_i} 确保复制资产的值始终为C_{t_i} 的值．注意，在给定t_i信息后，等式右边每一项都是已知的，因此，$\{\alpha_{t_i}, \beta_{t_i}\}$是非预报的．我们可以在这个“套期保值”(即复制资产组合)中应用鞅表示原理．

6.11.2 时间动态

考虑 C_{t_i} 在区间$[t,T]$上的变化，显然有：

$$C_T = C_t + \sum_{i=0}^{n} \Delta C_{t_i} \tag{6.120}$$

因为 $\Delta C_{t_i} = C_{t_{i+1}} - C_{t_i}$，表示成复制资产组合形式：

$$C_T = C_t + \sum_{i=0}^{n} \Delta[\alpha_{t_i} B_{t_i} + \beta_{t_i} S_{t_i}] \tag{6.121}$$

$$= C_t + \sum_{i=0}^{n} \Delta[\alpha_{t_i} B_{t_i}] + \sum_{i=0}^{n} \Delta[\beta_{t_i} S_{t_i}] \tag{6.122}$$

其中 Δ 表示一阶差分算子．

现在，回顾一下微分的“乘法法则”：

$$\mathrm{d}(uv) = v\mathrm{d}u + u\mathrm{d}v \tag{6.123}$$

代入(6.122)式右边第二项和第三项，[⊖]有

$$\sum_{i=0}^{n} \Delta[\alpha_{t_i} B_{t_i}] = \sum_{i=0}^{n} (\Delta\alpha_{t_i}) B_{t_{i+1}} + \sum_{i=1}^{n} \alpha_{t_i} (\Delta B_{t_i}) \tag{6.124}$$

$$\sum_{i=0}^{n} \Delta[\alpha_{t_i} S_{t_i}] = \sum_{i=0}^{n} (\Delta\alpha_{t_i}) B_{t_{i+1}} + \sum_{i=1}^{n} \alpha_{t_i} (\Delta B_{t_i}) \tag{6.125}$$

其中

$$\Delta[\alpha_{t_i} B_{t_i}] = [\alpha_{t_{i+1}} B_{t_{i+1}}] - [\alpha_{t_i} B_{t_i}]$$

$$\Delta\alpha_{t_i} = \alpha_{t_{i+1}} - \alpha_{t_i}$$

$$\Delta\beta_{t_i} = \beta_{t_{i+1}} - \beta_{t_i}$$

$$\Delta B_{t_i} = B_{t_{i+1}} - B_{t_i}$$

$$\Delta S_{t_i} = S_{t_{i+1}} - S_{t_i}$$

因此(6.122)可以写成：

$$C_T = C_t + \sum_{i=0}^{n} (\Delta\alpha_{t_i}) B_{t_{i+1}} + \sum_{i=1}^{n} \alpha_{t_i} (\Delta B_{t_i}) + \sum_{i=0}^{n} (\Delta\alpha_{t_i}) S_{t_{i+1}} + \sum_{i=1}^{n} \alpha_{t_i} (\Delta S_{t_i}) \tag{6.126}$$

⊖ 另一种得到下面等式的方法是简单代数运算．假设：$\Delta[\alpha_{t_i} B_{t_i}] = \alpha_{t_{i+1}} B_{t_{i+1}} - \alpha_{t_i} B_{t_i}$，注意我们可以在等是右边加上再减去 $\alpha_{t_i} B_{t_i}$，提公因式，然后就可以得到：

$$\alpha_{t_{i+1}} B_{t_{i+1}} - \alpha_{t_i} B_{t_i} = (\alpha_{t_{i+1}} - \alpha_{t_i}) B_{t_{i+1}} + \alpha_{t_i} (B_{t_{i+1}} - B_{t_i}) = \Delta\alpha_{t_i} B_{t_{i+1}} + \alpha_{t_i} \Delta B_{t_i}$$

重新组合后可以得到

106

$$C_T = C_t + \sum_{i=0}^{n}[(\Delta\alpha_{t_i})B_{t_{i+1}} + (\Delta\alpha_{t_i})S_{t_{i+1}}] + \sum_{i=0}^{n}[\alpha_{t_i}(\Delta S_{t_i}) + \alpha_{t_i}(\Delta S_{t_i})] \quad (6.127)$$

现在考虑表达式右边项，C_t是未知的，事实上，我们就是要找到确定这个无套利价格的方法，并使它满足(6.110)．带括号的另外两项需要详细讨论：

考虑带括号的第一项．根据t_{i+1}时刻的信息集，则括号内每一项都是已知的，$B_{t_{i+1}}$，$S_{t_{i+1}}$是市场价格，$\Delta\alpha_{t_i}$，$\Delta\beta_{t_i}$是金融分析师对复制资产组合权重的调整值．这里，带括号的第一项与鞅表示(6.109)中的D_t项有一些相似之处．

带括号的第二项在已知信息集I_{t_i}的情况下是未知的，因为它含有t_i后发生的价格变化值，可能需要信息集I_{t_i}所不包含的新信息才能确定．然而，尽管它们是未知的，一般来说这些价格变化都是可预测的，因此我们并不期望第二项起着鞅表示定理中dM_t的作用．另一方面，带括号的第二项一般有一个非零漂移项，因而不可能是鞅．

因此，这时我们不能对(6.127)使用期望算子$E_t^P[\cdot]$来得到类似

$$C_t = E_t^P[C_T]$$

的结果，其中 P 是真实世界概率．

一般情况下，(6.127)中带括号的项在经过这一运算后不会消失，但我们有两种工具可以选择．

第一种，我们可以将(6.127)中的$\{C_t, B_t, S_t\}$除以另一个无套利价格，并用标准化的价格而不是真实价格写出鞅表示．这种标准化如果做得好，可以保证C_t过程的任意漂移都被该漂移的标准化变量所“修正”，这种方法在我们想对未来收益C_T进行贴现时会非常方便．

第二种，当我们说第二个带括号项通常是可预测的而不是鞅时，我们是针对真实世界的概率而言的，我们可以用 Girsanov 定理将真实概率测度进行转化．也就是说，我们可以在风险中性概率下进行计算．⊖

现在我们来说明这些步骤是如何应用于(6.127)的．

6.11.3 标准化和风险中性概率

为了实施上面讨论的步骤，首先通过选择一个合适的价格将各项资产标准化．在这种情况下，除以B_t的价格使得标准化变得非常方便，定义

$$\widetilde{C}_t = \frac{C_t}{B_t}, \quad \widetilde{S}_t = \frac{S_t}{B_t}, \quad \widetilde{B}_t = \frac{B_t}{B_t} = 1 \quad (6.128)$$

注意到 $\widetilde{B}_t$是常数，不随时间增加，于是我们有

$$\Delta\widetilde{B}_t = 0, \quad 任意\ t_i \quad (6.129)$$

通过对B_t进行标准化显然可以消除该变量的趋势，但还不止这一点．

考虑标准化后在无穷小的区间 dt 内的预期变化．在连续时间下，

⊖ 我们将会在第 12 章和第 13 中详细讨论 Girsanov 定理．这里的讨论仅提供一些理论依据．

$$dB_t = rB_t dt \tag{6.130}$$

因为瞬时投资B_t的收益率等于无风险利率r. 代入

$$d\widetilde{S}_t = d\frac{S_t}{B_t} = \frac{dS_t}{B_t} - \widetilde{S}_t \frac{dB_t}{B_t} \tag{6.131}$$

$$= \frac{dS_t}{S_t}\widetilde{S}_t - \widetilde{S}_t r dt \tag{6.132}$$

其中我们用r代替dB_t/B_t.[⊖] 记得第二章中在无套利情况下通过对货币市场的标准化，S_t的预期收益率将会是无风险利率r：

$$E_t^Q\left[d\widetilde{S}_t\right] = E_t^Q\left[\frac{dS_t}{S_t}\widetilde{S}_t\right] - \widetilde{S}_r r dt \tag{6.133}$$

$$= r\widetilde{S}_t dt - \widetilde{S}_t r dt = 0 \tag{6.134}$$

其中Q是风险中性概率测度，如第2章中讨论的那样，它可以通过状态价格得到．因此标准化后的dS_t在Q测度下也具有零均值．107

同样，我们可以在离散时间下运用这种方法，来消除(6.127)中不想要的带括号的项. 由

$$\widetilde{C}_T = \widetilde{C}_t + \sum_{i=0}^{n}\left[(\Delta\alpha_{t_i})\widetilde{B}_{t_{i+1}} + (\Delta\alpha_{t_i})\widetilde{S}_{t_i}\right] + \sum_{i=0}^{n}\left[\alpha_{t_i}(\Delta\widetilde{S}_{t_i})\right] \tag{6.135}$$

在风险中性概率测度Q下，有

$$E_t^Q[\Delta\widetilde{S}_t] = 0 \tag{6.136}$$

因此对(6.135)进行期望运算$E_t^Q[\cdot]$可以得到

$$E_t^Q[\widetilde{C}_T] = \widetilde{C}_t + E_t^Q\left\{\sum_{i=0}^{n}\left[\left(\Delta\alpha_{t_i}\right)\widetilde{B}_{t_{i+1}} + \left(\Delta\alpha_{t_i}\right)\widetilde{S}_{t_i}\right]\right\} + E_t^Q\left[\sum_{i=0}^{n}\left[\alpha_{t_i}\left(\Delta\widetilde{S}_{t_i}\right)\right]\right] \tag{6.137}$$

$$= \widetilde{C}_t + E_t^Q\left\{\sum_{i=0}^{n}\left[\left(\Delta\alpha_{t_i}\right)\widetilde{B}_{t_{i+1}} + \left(\Delta\alpha_{t_i}\right)\widetilde{S}_{t_i}\right]\right\} + 0 \tag{6.138}$$

显然，如果我们能够消除带括号的项，我们就能得到想要的结果

$$C_t = B_t E_t^Q\left[\frac{C_T}{B_T}\right] \tag{6.139}$$

即未知量C_t的无套利价格．

那么如何消除(6.136)中后面一项带括号项呢？我们通过选择适当的$\{\alpha_{t_i}, \beta_{t_i}\}$达到这一目的，对于

$$E_t^Q\left\{\sum_{i=0}^{n}\left[\left(\Delta\alpha_{t_i}\right)\widetilde{B}_{t_{i+1}} + \left(\Delta\alpha_{t_i}\right)\widetilde{S}_{t_i}\right]\right\} \tag{6.140}$$

其中的$\{\alpha_{t_i}, \beta_{t_i}\}$能够确保复制资产组合是自融资的．即对任意$i$，只要满足

$$\alpha_{t_{i-1}} B_{t_i} + \beta_{t_{i-1}} S_{t_i} = \alpha_{t_i} B_{t_i} + \beta_{t_i} S_{t_i} \tag{6.141}$$

⊖ 因为B_t是确定的，而S_t具有线性型，这里没有伊藤修正项．

即可．也就是说，t_i时刻选择的资产组合在t_{i+1}时刻的价值足以对资产组合权重进行重新调整．注意，(6.141)是在非标准化价格下的表示，因为无论我们使用哪种标准化，它都会在等号两边消掉．

6.11.4 总结

现在我们可以从资产定价的角度对这些计算进行总结．

前面几节中的计算基本上是基于 3 种工具．第一种是鞅表示定理，它说明给定包含趋势随机过程，我们可以将其分解成一个已知的趋势和一个鞅．尽管这一结果看起来技巧性很强，但实际上它非常直观．给定时间序列，原则上可以将它的趋势和趋势附近的波动分开．市场参与者接触现实数据时，都默认这种规律存在，事实上这就是在使用鞅表示定理的粗糙形式．

我们使用的第二种工具是标准化，鞅表示定理应用于标准化的价格而不是观察到的价格．这消除了鞅表示中一些我们不希望出现的项．

108

第三种工具是测度变换，通过利用风险中性概率计算期望，我们可以确保鞅表示中仍存在的不想要的项会被去掉．事实上，使用风险中性测度对于改变S_t过程的预期趋势是有影响的，标准化确保了新的趋势随着B_t的增长被消掉了．结果就是标准化的C_t最终没有任何趋势，变成了鞅．如果使用自融资复制资产组合方法，就得到了定价公式(6.127).

6.12 小结

本章主要介绍了鞅工具．鞅是一种没有时间趋势的随机过程，我们介绍的几个例子，它们在后面的章节中将会非常有用．

本章还介绍了将具有正(或负)的时间趋势的随机过程化为鞅的方法．

最后我们介绍一下这些理论概念与从业者的关联．

用 S_t表示交易者在 t 时刻观察到的资产价格．在无穷小的时间区间内，交易者获得了关于S_t的无法预测的新信息．我们将其表示为$dS_t=\sigma_t dW_t$，其中σ_t表示波动率，dW_t是布朗运动的增量，波动率有时间下标，表示随时间变化．还要注意 dS_t没有可预测的漂移项．

在更长的区间内，这种不可预测的信息会积累，在间隔 T 后，资产价格变为 $S_{t+T}=S_t+\int_t^{t+T}\sigma_u dW_u$.

该等式与(6.106)具有相同形式．如果每一个增量信息都是不可预测的，那么增量信息的和也应该是不可预测的．但这意味着 S_t应该是鞅，我们必须有 $E_t\left[\int_t^{t+T}\sigma_u dW_u\right]=0$.

这是随机积分的一个重要性质．但是在市场上，信息流也为金融市场参与者施加了限制条件，鞅方法是讨论这些等式的中心，因而它们对于从业者来说是不可缺少的．

6.13 参考阅读

如果读者希望了解更多关于鞅理论的知识，可以参考 Williams(1991)的介绍性读本，这本书的可读性非常强，并且利用简单的模型对所有主要的鞅结果进行了详细的讨论. Revuz 和 Yor(1994)是鞅方面的一本优秀的高级教材. Shiryayev(1984)的综述对这些结果进行了中间层次的处理. 对于交易收益和随机积分，读者可以参考 Cox 和 Huang(1989). Dellacherie 和 Meyer(1980)对于鞅进行了全面的介绍. Musiela 和 Rutkowski(1997)是一本在资产定价的鞅方法方面非常优秀和全面的书.

6.14 习题

1. Y 是随机变量，且

$$E[Y]<\infty$$

(a) 证明M_t是鞅，其中 $M_t=E[Y\mid I_t]$

109

(b) 给定递增信息集

$$\{I_0 \subseteq \cdots \subseteq I_t \subseteq I_{t+1} \subseteq \cdots\}$$

这意味着所有条件期望都是鞅吗?

2. 考虑随机变量:

$$X_n = \sum_{i=1}^{n} B_i$$

其中每一个B_i都是抛硬币的结果: $B_i=\begin{cases}+1 & \text{正面朝上}\\ -1 & \text{背面朝上}\end{cases}$

令 $n=4$，考虑X_4.

(a) 计算 $E[X_4\mid I_1]$，$E[X_4\mid I_2]$，$E[X_4\mid I_4]$

(b) 令 $Z_i=E[X_4\mid I_i]$，Z_i，$i=1$，…，4 是鞅吗?

(c) 定义: $V_i=B_i+\sqrt{i}$，$\widetilde{X}_n = \sum_{i=1}^{n} V_i$，$V_i$ 是鞅吗?

(d) 可以通过适当的变换将V_i转化为鞅吗?

(e) 可以通过改变抛硬币的概率将V_i转化为鞅吗?

3. W_t是维纳过程，t 表示时间. 下面的随机过程是鞅吗?

(a) $X_t=2W_t+t$ (b) $X_t=W_t^2$ (c) $X_t = W_t^2 - 2\int_0^t sW_s\,ds$

4. 假设有 $M_T(X_t) = M_0(X_0) + \int_0^T g(t,X_t)dW_t$，等式在给定信息集序列$\{I_t\}$下成立. 标的过程$X_t$服从 SDE:

$$dX_t = \mu dt + \sigma dW_t$$

确定上面表示中的 $g(\cdot)$在不同的M_t下是否为鞅:

(a) $M_T(X_t)=W_T$. (b) $M_T(X_T)=W_T^2$. (c) $M_T(X_T)=e^{W_T}$.

5. 给定如下表示:

$$M_T(X_t) = M_0(X_0) + \int_0^T g(t, X_t) \mathrm{d}W_t$$

如果$M_T(X_T)$是普通欧式看涨期权到期时的收益，$g(\cdot)$可以被确定吗？即，如果$M_T(X_T)$满足

$$M_T(X_T) = \max[X_T - K, 0]$$

其中 $0<K<\infty$是执行价格．困难在哪里？

6. X 是随机过程，且对于任意 $t\geqslant 0$，都有 $E[X_t]=0$，并且 X 是独立增量过程(即X_t-X_s与X_s相互独立，对于任意 s，t，$0\leqslant s<t<\infty$).

(a) 证明 X 是鞅．

(b) 假设 X 具有独立增量，但 $t\to E[X_t]$，$E[X_t]$不是常数．X 还是鞅吗？证明．

110

7. 考虑随机过程$X_t=W_t^3$．通过模拟W_t的值计算其数学期望，考察该过程是否为鞅．

第7章　随机环境下的微分

7.1　引言

第3章复习了确定性环境下的微分，函数$f(x)$关于x的导数刻画了$f(\cdot)$在x发生细微变化(记为dx)时的变化率．而$f(\cdot)$的变化可以通过下列公式计算：

$$df = f_x dx \tag{7.1}$$

其中f_x表示$f(x)$关于x的导数．

在随机环境下我们也需要类似的概念．例如，已知标的资产价格S_t变化时，以S_t为标的的看涨期权价格将会如何变化？在确定性环境下，我们用“普通”的微分法则去处理这类问题. 然而在金融资产定价领域，我们涉及的是随机的变量，且风险的概念起着核心作用. 那么当标的变量是连续时间的随机过程时，也可以使用类似的公式吗？

微分的概念与常微分方程(ODE)紧密相联，在常微分方程中，某一变量的变化对其他变量的影响可以用方程确切地表示出来．事实上，(向量)微分方程是对确定性过程中的动态变化进行建模的规范方法，当然在此过程中，我们需要相关的导数存在．

微分方程也可以用于对资产价格的动态变化进行建模吗？要对资产价格进行建模，首先要面对的困难就是资产价格的随机性．热量在金属棒中的传递方式可以用确定性的模型进行合理地近似，但是，在对衍生资产进行定价时，标的工具的随机性是必不可少的，因为正是消除或承担风险的需求催生了衍生资产．在确定性环境中，所有情况都完全可以被预测，不存在风险，因此也不需要金融衍生品．如果随机性是必不可少的，那么随机环境下的微分将如何定义呢？

可以简单地在常微分方程后添加一个随机误差项来对金融衍生品进行定价吗？或者说，在定义随机微分方程(SDE)时还存在什么其他的困难吗？

本章使用随机微分方程这一基本模型，来解决随机环境下的微分问题．我们将从头开始构建SDE，然后说明直接将确定性微积分中的微分公式用于随机环境将会面临的困难．更准确地说，我们将首先说明在什么情况下，连续时间过程S_t的变化可以用随机微分方程

$$dS_t = a(S_t, t)dt + b(S_t, t)dW_t \tag{7.2}$$

所描述的动态过程进行模拟，其中dW_t是干扰项，它表示在无穷小的区间dt上发生的不可预测的事件．$a(S_t, t)$、$b(S_t, t)$分别是漂移系数和扩散系数，它们都是I_t适应的．然后，我们将讨论干扰项dW_t的性质，因为它驱动了整个系统，是潜在随机性的来源．我们将看到W_t是一个非常不规则的过程，它的导数在确定性微积分意义下不存在．因此，dS_t、dW_t这种增量需要通过其他方式进行定义．

从头开始构建随机微分方程还有一个附带的好处，通过这个过程，我们可以熟悉连续时间下的随机微积分．它将连续时间下的积分与离散时间下的求和的计算联系到了一起，

并且可以消除一些错误的观念.

7.2 问题起源

本节将对确定性环境和随机性环境下的微分进行具体的比较.

用 S_t 表示证券价格, $F(S_t, t)$ 表示以 S_t 为标的资产的衍生品价格. 股票经纪人感兴趣的是证券价格在下一瞬间的增量 $\mathrm{d}S_t$. 另一方面, 衍生品交易柜台需要知道的是 $\mathrm{d}F_t$, 以 S_t 为标的资产的衍生品价格的变化. 如何从对 $\mathrm{d}S_t$ 的估计出发, 计算 $\mathrm{d}F_t$ 呢?

在这里, 我们感兴趣的不是标的证券的价格如何变化, 而是金融衍生品价格对标的资产价格的变化做出怎样的反应. 也就是说, 需要使用"链式法则". 如果这里可以用普通的微积分, 市场参与者就可以用使用公式

$$\mathrm{d}F_t = \frac{\partial F}{\partial S}\mathrm{d}S_t \tag{7.3}$$

或者, 写成偏导数形式

$$\mathrm{d}F_t = F_s \mathrm{d}S_t \tag{7.4}$$

但是确定性微积分的法则真的适用于这种情况吗? 链式法则也可以用于随机环境下吗?

下面我们将说明随机环境下的微分法则是不同的. 我们继续用 x 的函数 $f(x)$ 进行讨论. 如第 3 章所讨论, 普通的微分是用极限运算

112

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = f_x \tag{7.5}$$

来定义的, 其中该极限要求满足

$$f_x < \infty \tag{7.6}$$

这里, $f(x+h) - f(x)$ 表示当 x 有增量 h 时, 函数值发生的相应变化. 因此, 如果 x 表示时间, 那么 $f(x)$ 的导数就是 $f(x)$ 在无穷小的时间区间内的变化率.[㊀] 在这种情况下, 时间是确定性变量, 可以使用"普通"微积分.

但是, 如果 $f(x)$ 中的 x 是连续时间轴上的一个随机变量呢? 可以用类似的方式定义导数, 并且使用常规的链式法则吗? 一般来说, 这个问题的答案是否定的. 我们用一个具有启发性意义的讨论来说明这个问题.

假设 $f(x)$ 是随机过程 x 的函数.[㊁] 在 x 的某个已知值 x_0 附近对 $f(x)$ 进行展开.[㊂] 其泰勒展开式为

$$f(x) = f(x_0) + f_x(x_0)(x - x_0) + \frac{1}{2} f_{xx}(x_0)(x - x_0)^2 + \\ + \frac{1}{3!} f_{xxx}(x_0)(x - x_0)^3 + R(x, x_0) \tag{7.7}$$

㊀ 用 $f(x+h) - f(x)$ 除以 h, 我们就得到了一个比率. 这个比率告诉我们 $f(x)$ 在每单位 h 内发生的变化. 因此, 该导数是变化速度.

㊁ 为了记号的简便, 我们省略了 x 的时间下标.

㊂ 感兴趣的读者可以回顾第 3 章, 复习泰勒级数展开.

其中 $R(x,x_0)$ 表示泰勒展开式的余项．注意，该余项包含：$f(x)$ 高于 3 阶的偏导数，高于 3 阶的阶乘及 $(x-x_0)$ 高于 3 阶的幂次项．

现在我们来分析泰勒展开式的近似，考虑等号右边除 $R(x,x_0)$ 外的项．

如果令

$$\Delta x = x - x_0 \tag{7.8}$$

那么 $f(x)$ 可写为 $f(x_0+\Delta x)$. 泰勒展开式的近似项变为[⊖]

$$f(x_0+\Delta x)-f(x_0)\approx f_x(\Delta x)+\frac{1}{2}f_{xx}(\Delta x)^2+\frac{1}{6}f_{xxx}(\Delta x)^3 \tag{7.9}$$

在该表达式右边，Δx 表示随机变量 x 发生的"小"变化．注意，尽管我们认为这一变化很小，但是我们并不希望它小到可以忽略不计．毕竟，我们的目的是确定 x 的变化给 $f(x)$ 带来的影响，如果认为 x 的变化可以忽略不计，那么它造成的影响将无法确定．因此，在对右边式子进行近似处理时，我们希望保留 $f_x(\Delta x)$ 项．

现考虑右边第二项．如果变量 x 是确定性的，我们可以认为 $(\Delta x)^2$ 非常小．虽然 Δx 很小但不可忽略，而 $(\Delta x)^2$ 足够小以致可以被忽略．事实上，如果 Δx 非常小，它的平方会更小，所以在某种情况下是可以被忽略的．但是，在本例中，x 是随机变量，因此 x 的变化也是随机的. 假设这些变化的均值为 0，这些随机变量仍然是随机取值的，因为它有一个正的方差：

$$E[\Delta x]^2 > 0 \tag{7.10}$$

但是从字面上看，该式意味着"平均来说"，$(\Delta x)^2$ 是非零的．也就是说，只要 x 是随机变量，令 $(\Delta x)^2$ 等于 0 就等价于令 Δx 的方差为 0. 这意味着，用非随机量来对随机变量 x 进行近似产生的效果将违背我们的初衷．毕竟，我们在努力探究的是 x 的随机变化对 $f(x)$ 的影响．因此，只要 x 是随机的，右边的泰勒近似就必须保留二阶项．

113

另一方面，尽管需要保留右边式子中 Δx 的一阶和二阶项，舍去 Δx 的三阶及以上项仍然是合理的．即高阶项忽略不计，也不会造成太大的麻烦[⊜].

因此，泰勒近似的一种备选形式为

$$f(x_0+\Delta x)-f(x_0)\approx f_x(\Delta x)+\frac{1}{2}f_{xx}E[(\Delta x)^2] \tag{7.11}$$

其中 $(\Delta x)^2$ 换成了它的期望值，这种近似方法相当于用某项的"平均"值来替代它．在本书的第二部分，我们将引入处理这类问题要用到的工具．

替代使用 $E[(\Delta x)^2]$，另一种可能性是在时间区间长度趋于 0 时使用随机变量 $(\Delta x)^2$ 的某种极限，我们在第 4 章中已经讨论过这种近似．可以证明，在相同的条件下，这两种方法得到的结果是相同的．事实上，如果 h 代表时间区间，在这一区间内 x 发生的变化为 Δx，并且 h"很小"，那么在均方意义下 $(\Delta x)^2$ 将趋近于 $\sigma^2 h$.

⊖ 在后面为了记法的简便，我们省略了 $f_x(x_0)$，$f_{xx}(x_0)$，$f_{xxx}(x_0)$ 中的自变量．

⊜ 一些读者可能记得第 6 章中关于连续时间鞅的讨论．我们在那里证明了：对于连续平方可积鞅，一次变差无界，二次变差收敛于一个有定义的随机变量，高次变差不存在．因此，如果 x 是连续平方可积鞅，在某些近似意义下，可以认为 Δx 的高阶项等于 0.

因此，我们有两种可能的近似式，这取决于 x 是否是随机变量．

如果 x 是随机的，有

$$f(x_0+\Delta x)-f(x_0)\approx f_x(\Delta x)+\frac{1}{2}f_{xx}E[(\Delta x)^2] \tag{7.12}$$

或

$$f(x_0+\Delta x)-f(x_0)\approx f_x(\Delta x)+\frac{1}{2}f_{xx}E[x^*] \tag{7.13}$$

其中 x^* 是 $(\Delta x)^2$ 的均方极限．

一旦 x 变为确定性的，我们就可以假定当 Δx 很小时 $(\Delta x)^2$ 可以忽略不计，并有

$$f(x_0+\Delta x)-f(x_0)\approx f_x(\Delta x) \tag{7.14}$$

上述结果为处理确定性和随机性环境下的微分提供了方法．

比如，如果有(7.14)式，我们可以在两边同时除以 Δx，得到近似式

$$\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}\approx f_x \tag{7.15}$$

但是对于随机的 Δx，我们仍不清楚是否可以忽略第三项，令(7.8)中的 $\Delta x\to 0$，

$$\lim_{\Delta x\to 0}\frac{f(x+\Delta x)-f(x)}{\Delta x}\approx f_x+\frac{1}{2}f_{xx}\lim_{\Delta x\to 0}\frac{(\Delta x)^2}{\Delta x}$$

这就定义了导数，接下来我们就对这个式子进行讨论．

7.3 一个讨论微分的框架

微分概念讨论的是无穷小区间上发生的增量变化．将它应用到金融市场上，我们可以研究资产价格随时间发生的变化，当然，我们假定这些变化是随机的．因此，在随机微积分中，导数的定义需要用到某种类型的概率收敛．⊖

114

讨论随机微分，一个与之联系最密切的应用是随机微分方程(SDE)：

$$\mathrm{d}(S_t,t)=a(S_t,t)\mathrm{d}t+b(S_t,t)\mathrm{d}W_t \tag{7.16}$$

为了理解如何在随机环境下进行微分，我们将从头开始“构建”SDE. 先从离散时间开始，然后扩展到连续时间上去．

考虑时间区间 $t\in[0,T]$. 考虑图 7-1，

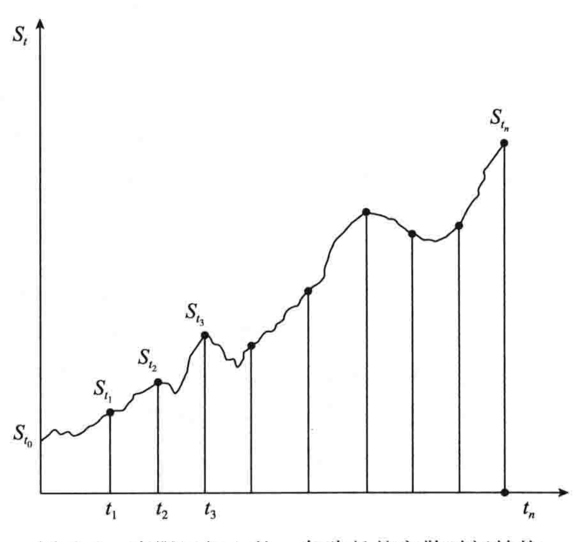


图 7-1 有限区间上的一条路径的离散时间结构

⊖ 在概率收敛中，我们感兴趣的是找到一列随机变量 x_n 所收敛的那个随机变量 x^*．对于“大的”n 来说，极限随机变量 x^* 可以用于对 x_n 进行近似，因为极限随机变量一般比 x_n 本身更容易处理．

将 x 轴上 $[0,T]$ 区间平均分成 n 个长度为 h 的小区间．仍用前面几章中使用的符号，考虑如下分割：

$$0 = t_0 < t_1 < \cdots < t_n = T \tag{7.17}$$

本章的主要不同在于，对任意 k 都有

$$t_k - t_{k-1} = h \tag{7.18}$$

这意味着

$$t_k = kh \tag{7.19}$$

115

因此，我们有如下关系

$$n = \frac{T}{h} \tag{7.20}$$

定义在这些有限区间内观察到的变量：

$$S_k = S(kh) \tag{7.21}$$

及证券价格 $S(t)$ 在有限区间 h 内发生的变化

$$\Delta S_k = S(kh) - S((k-1)h) \tag{7.22}$$

现在，选定一个特定的区间 k. 只要相应的数学期望存在，我们总可以用如下方式定义随机变量 ΔW_k：

$$\Delta W_k = [S_k - S_{k-1}] - E_{k-1}[S_k - S_{k-1}] \tag{7.23}$$

这里，符号 $E_{k-1}[\cdot]$ 表示以第 $k-1$ 个区间结束时所有可获得的信息为条件的条件数学期望. 已知第 $(k-1)$ 个区间结束时所有信息，ΔW_k 是 $[S_k - S_{k-1}]$ 中完全无法预测的部分．等式右边的第一项表示资产价格 $S(t)$ 在第 k 个区间内发生的实际变化，第二项是市场参与者在已知信息集 I_{k-1} 时，所预测的变化[⊖]. 我们称不可预测的部分为“干扰”.

该干扰项有以下性质．

- 当第 $(k-1)$ 个区间结束时，ΔW_k 是未知的，它在第 k 个区间结束时才可以被观测到．按测度论中的术语，我们称 ΔW_k 是 I_k 可测的，即给定集合 I_k，可以得到 ΔW_k 的精确值．
- 已知 $k-1$ 时刻的信息集时，ΔW_k 的值是不可预测的，且

$$E[\Delta W_k] = 0, \quad 任意\ k \tag{7.24}$$

- ΔW_k 是鞅过程的变化，称为鞅差分．累积误差 W_k 为

$$W_k = \Delta W_1 + \cdots + \Delta W_k \tag{7.25}$$

$$= \sum_{i=1}^{k} \Delta W_i \tag{7.26}$$

其中，我们假设 W_0 的初始值为 0.

W_k 是鞅，是因为：

⊖ 如果已知信息集完全无法为确定 $S(t)$ 在未来的变化提供信息，那么该预测值为 0. 在这种情况下，$[S_k - S_{k-1}]$ 本身就是不可预测的部分．

$$E_{k-1}[W_k] = E_{k-1}[\Delta W_1 + \cdots + \Delta W_k] = \Delta W_1 + \cdots \Delta W_{k-1} + E_{k-1}[\Delta W_k] \quad (7.27)$$
$$= W_{k-1} \quad (7.28)$$

最后一个等式成立是因为给定I_{k-1}时，$E_{k-1}[\Delta W_k]$等于0，且$\Delta W_i(i=1, \cdots, k-1)$是已知的．

像ΔW_k这样的随机变量，它的重要性在哪里呢?

考虑金融市场参与者．对于决策者来说，资产价格中包含的重要信息就是ΔW_k．这些不可预测的"消息"持续发生，并且可以通过路透、彭博等主要网络"实时"观测到．因此，资产价格的"实时"变动是由ΔW_k所主导的．这意味着，为了讨论随机环境下的微分，我们需要研究ΔW_k的性质．特别地，我们将证明，在某些可被接受的假设条件下，ΔW_k^2及它的无穷小等价形式dW_k^2在泰勒近似中都是不可以"忽略的"．

7.4 增量误差的度量

干扰项ΔW_k代表无法预测的变化，ΔW_k^2是它的平方．在确定性环境中，微分的概念涉及的是ΔW_k等项，并认为其平方项可以忽略不计．实际上在确定性微积分中，像ΔW_k^2这种项在微分过程中并根本就不会出现[⊖]．然而在随机微分中，通常需要考虑二阶项的变化，本节将讨论这些项的近似．

116

有两种方法可以对其进行近似，一种是随机过程中使用的方法，另一种方法是Merton(1990)中所讨论的方法．这里我们使用Merton的方法，因为这样更容易理解后续假定背后的经济学思想．Merton方法是通过研究金融市场上信息流的特征，并且通过某种精确的方式对其进行建模的方法．

首先，我们先定义一些记号．

将ΔW_k的(无条件)方差记为V^k：

$$V^k = E_0[\Delta W_k^2] \quad (7.29)$$

定义累积误差项的方差为：

$$V = E_0\Big[\sum_{k=1}^{n} \Delta W_k\Big]^2 = \sum_{k=1}^{n} V^k \quad (7.30)$$

这里用到了ΔW_k关于k不相关，交叉乘积项的期望为0的性质．

现在我们按照Merton(1990)的思路来引入一些假定．

假定1

$$V > A_1 > 0 \quad (7.31)$$

其中A_1独立于n.

该假定意味着证券价格的波动率存在一个下界．这说明当区间$[0,T]$被划分成越来越

⊖ 更高阶的导数与此相同．

细的子区间时，[1]

$$n \to \infty \tag{7.32}$$

累积误差项的方差 V 大于 0. 也就是说，对证券价格越来越频繁的观测不会消除所有的“风险”. 显然，大多数金融市场参与者都会接受这一假定. 即使我们在越来越短的时间区间内观测市场，资产价格的不确定性也不会消失.

假定 2

$$V < A_2 < \infty \tag{7.33}$$

其中A_2独立于 n.

该假定意味着，累积误差项的方差存在一个上界，使波动率被限制在该上界以下. 随着时间轴被分割成越来越小的子区间，我们可以进行越来越频繁的交易. 这种交易不会给整个系统带来无限的不稳定性. 大部分市场参与者也会认同这一假设. 毕竟，允许越来越频繁的交易和在线观测不会产生无穷大的波动率.

为了给出第 3 个假定，定义

$$V_{\max} = \max_k \left[V^k, k = 1, \cdots, n \right] \tag{7.34}$$

即$V_{\max}$是资产价格在波动性最强的子区间内的方差.

假定 3

$$\frac{V^k}{V_{\max}} > A_3, \quad 0 < A_3 < 1 \tag{7.35}$$

其中A_3独立于 n.

根据这一假定，金融市场的不确定性不会集中于某个特定时期，不论市场在何时开放，都一定会存在某种波动. 这一假设排除了金融市场上彩票式的不确定性.

117

现在我们可以开始讨论$(\Delta W_k)^2$的一条非常重要的性质了.

下述命题是随机微积分的核心结论.

命题 1 在假定 1、2、3 下，ΔW_k的方差与 h 成正比，

$$E[\Delta W_k]^2 = \sigma_k^2 h \tag{7.36}$$

其中 σ_k 是不依赖于 h 的有限常数，但可能依赖于在 $k-1$ 时刻的信息.

根据该命题，当 h 变得越来越小时，资产价格波动也变得越来越小.

由于该命题是核心结论，下面我们给出其简要的证明.

证明：根据假定 3：

$$V^k > A_3 V_{\max} \tag{7.37}$$

两边同时对所有区间求和：

$$\sum_{k=1}^{n} V^k > n A_3 V_{\max} \tag{7.38}$$

---

[1] 子区间有着相同的长度 h.

假定 2 说明上式左边存在上界：

$$A_2 > \sum_{k=1}^{n} V^k > nA_3 V_{\max} \tag{7.39}$$

两边同时除以 nA_3：

$$\frac{1}{n}\frac{A_2}{A_3} > V_{\max} \tag{7.40}$$

注意到 $n=\frac{T}{h}$，则有

$$\frac{1}{n}\frac{A_2}{A_3} > V_{\max} > V^k \tag{7.41}$$

$$\frac{h}{T}\frac{A_2}{A_3} > V^k \tag{7.42}$$

上式给出了V^k只与 h 有关的上界．现在我们要得到一个只与 h 有关的下界．已知

$$\sum_{k=1}^{n} V^k > A_1 \tag{7.43}$$

则有

$$nV_{\max} > \sum_{k=1}^{n} V^k > A_1 \tag{7.44}$$

将(7.44)两边同时除以 n

$$V_{\max} > \frac{A_1}{n} \tag{7.45}$$

$$V_{\max} > \frac{A_1}{T}h \tag{7.46}$$

根据假定 3

$$V^k > A_3 V_{\max} > \frac{A_3 A_1}{T}h \tag{7.47}$$

即

$$V^k > \frac{A_3 A_1}{T}h \tag{7.48}$$

因此有

$$\frac{h}{T}\frac{A_2}{A_3} > V^k > \frac{A_3 A_1}{T}h \tag{7.49}$$

显然，无论 n 等于多少，方差项V^k都同时具有与 h 成正比的上界和下界．这意味着我们可以找到一个依赖于 k 的常数σ_k，在忽略 h(更小的)的高阶项时，使V^k与 h 成正比：

$$V^k = E[\Delta W_k]^2 = \sigma_k^2 h \tag{7.50}$$

7.5 命题 1 的隐含结论

命题 1 有几个隐含结论，其中之一为：首先，如果相关数学期望存在，则总有

$$S_k - S_{k-1} = E_{k-1}[S_k - S_{k-1}] + \sigma_k \Delta W_k \tag{7.51}$$

118

其中 ΔW_k的方差为 h⊖，两边同时除以 h：

$$\frac{S_k - S_{k-1}}{h} = \frac{E_{k-1}[S_k - S_{k-1}]}{h} + \frac{\sigma_k \Delta W_k}{h} \tag{7.52}$$

另一方面，根据上述命题

$$E[\Delta W_k]^2 = h \tag{7.53}$$

假设我们取 h 作为 ΔW_k^2的近似：

$$\Delta W_k^2 \approx h \tag{7.54}$$

（在第 9 章中，我们将证明这种近似在均方收敛意义下是有效的.）

在第 3 章中对导数的概念进行标准定义时，我们令 h 趋于 0. 在这里也使用相同的做法，并假设我们可以得到这一随机变量的“极限”：

$$\lim_{h \to 0} \frac{W_{(k-1)h+h} - W_{(k-1)h}}{h} \tag{7.55}$$

那么可以认为它是W_t关于时间 t 的导数. 而(7.54)中的近似说明这一导数可能没有定义：

$$\lim_{h \to 0} \frac{|W_{(k-1)h+h} - W_{(k-1)h}|}{h} \to \infty$$

图 7-2 从几何图形上刻画了这一点. 画出函数 $f(h)$：$\frac{h^{1/2}}{h}$的图像，显然，当 h 越来越小时，$f(h)$趋于无穷大，不存在有限的极限.

当然，这里出现的矛盾非常具有启发性. 当极限运算是作用在随机变量上而不是确定性函数上时，如何进行规范化仍不得而知. 但是这一矛盾仍然非常具有建设性，因为它说明了无穷小区间上不可预测的“消息”的基本特征，即

$$E[\sigma_k \Delta W_k]^2 = \sigma_k^2 h$$

而这会为定义随机环境下关于时间的导数带来难以克服的困难.

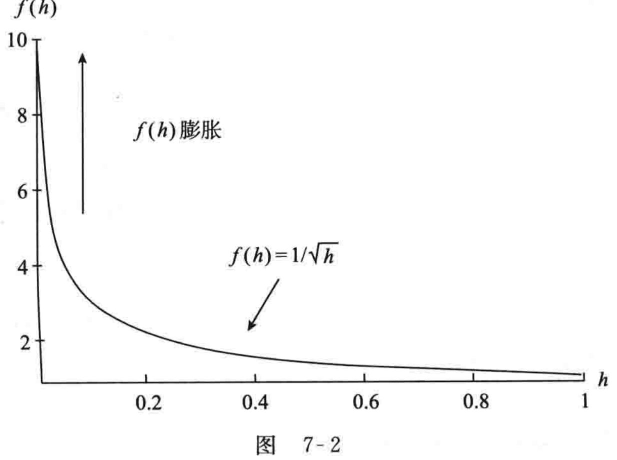


图 7-2

7.6 归并结果

至此，我们完成了两件事. 首先，任意随机过程S_t都可以将其在某个有限区间 h 上的变化写成

$$S_k - S_{k-1} = E_{k-1}[S_k - S_{k-1}] + \sigma_k \Delta W_k \tag{7.56}$$

⊖ 在该等式中，参数 σ_k 是 ΔW_k 的系数. 这是一个平凡变换，因为 $\sigma_k \Delta W_k$ 的方差现在等于 $\sigma^2 h$.

其中 ΔW_k 项在该时间区间开始时的信息下是无法观测的．㊀

119

其次，我们说明了，当 h 很小时，无法观测的干扰项具有与时间区间的长度 h 成正比的方差：

$$V[\Delta W_k] = h \tag{7.57}$$

为了得到定义在有限区间上的微分方程，我们还需要下面的第三步和第四步．我们对(7.56)式右边的第一项进行近似

$$E_{k-1}(S_k - S_{k-1}) \tag{7.58}$$

这一项是条件期望，或者说是对资产价格变化的预测．它的大小取决于最新的信息集和所考虑时间区间的长度．因此，$E_{k-1}(S_k-S_{k-1})$可以被写作

$$E_{k-1}(S_k - S_{k-1}) = A(I_{k-1},h) \tag{7.59}$$

其中 $A(\cdot)$是某个函数．从这个角度来看显然有：如果 $A(\cdot)$是关于 h 的光滑函数，那么它在 $h=0$ 附近有泰勒展开式

$$A(I_{k-1},h) = A(I_{k-1},0) + a(I_{k-1})h + R(I_{k-1},h) \tag{7.60}$$

其中，$a(I_{k-1})$是 $A(I_{k-1},h)$在 $h=0$ 处关于 h 的一阶导数，$R(I_{k-1},h)$是该泰勒展开式的余项㊁．如果 $h=0$，时间不发生变化，那么资产价格变化的预测值也为 0．也就是说，

$$A(I_{k-1},0) = 0 \tag{7.61}$$

与确定性常微分方程一样，处理随机常微分方程的惯例是，任何阶数高于 1 的 h 的幂次项都小到忽略不计．㊂

因此，与普通微积分类似，可以令

$$R(I_{k-1},0) \approx 0 \tag{7.62}$$

从而得到一阶的泰勒展开近似：

$$E_{k-1}(S_k - S_{k-1}) \approx a(I_{k-1},kh)h \tag{7.63}$$

利用这些结果，就可以将(7.56)式重新写成随机微分方程的形式：㊃

$$S_{kh} - S_{(k-1)h} \approx a(I_{k-1},kh)h + \sigma_k\left[W_{kh} - W_{(k-1)h}\right] \tag{7.64}$$

在后面的各章中，我们会令 $h\rightarrow 0$，进而得到(7.56)式在无穷小区间上的形式，即随机微分方程(SDE)：

$$\mathrm{d}S_t = a(I_t,t)\mathrm{d}t + \sigma_t \mathrm{d}W_t \tag{7.65}$$

我们称 $a(I_t,t)$为该随机微分方程的漂移项，σ_t为扩散项．

随机微分

我们在本章中多处提到了随机增量的极限．显然，我们需要得出对于如 $\mathrm{d}S_t$，$\mathrm{d}W_t$等增

㊀ 假设相应的数学期望存在．

㊁ 在 I_{k-1}给定的条件下，我们涉及的是非随机的数量关系，泰勒展开中的导数可以按照标准方式定义．

㊂ 由于 h^2 是确定性函数，因此这与标准微积分中忽略微分方程中所有二阶项的做法一致．

㊃ 这里在 S_k 和 W_k 的记号中再次引入了 h，说明这些项是依赖于 h 的．

量的规范定义．

如何使这些项变得更明确呢？

为此，我们需要对伊藤积分的基本概念进行定义．只有借助伊藤积分，我们才能对如 dS_t，dW_t 等量进行规范定义，从而对于随机微分方程中涉及的工具进行充分说明．我们将在第 9 章中介绍伊藤积分．

7.7 小结

120

普通微积分中的微分不能直接推广为随机导数，因为随机过程在无穷小的区间上变化的方差不等于 0. 此外，当新的信息流服从某种合适的假定时，连续时间随机过程会变得不稳定，关于时间的导数可能不存在．在小区间内，ΔW_k 由 h 决定，当 h 越来越小时，ΔW_k 与 h 之比的绝对值可能越来越大，极限可能没有定义．

另一方面，尽管定义微分存在一些困难，我们需要在一定的假设条件下才能构建 SDE. 在这一意义上，随机微分方程十分具有代表性，它可以适用于一大类随机过程．通过将随机过程的变化分解成可观测和不可观测的两部分，并给出一些关于可观测部分光滑性的假设，就可以构造出随机微分方程．

7.8 参考阅读

无法观测的误差在 3 个假定下，方差与 h 成正比的证明来自 Merton(1990). Merton(1990)中关于连续时间的数理金融可能对理解本章有所帮助．

7.9 习题

1. 考虑在 Black-Scholes 分析中具有基础作用的随机过程 S_t：

$$S_t = S_0 e^{\mu t + \sigma W_t}$$

其中 W_t 是维纳过程，且 $W_0=0$，μ 是"趋势"因子，且

$$(W_t - W_s) \sim \mathcal{N}(0,(t-s))$$

也就是说 W_t 的增量均值为 0，方差为 $t-s$. 因此，W_t 的增量在 t 时刻的方差等于自观测到 W_s 起所流逝的时间．我们还知道这些维纳增量具有时间独立性．据此，可以认为 S_t 是服从对数正态分布的随机变量．我们希望对这个过程可能的路径进行讨论．令 $\mu=0.01$，$\sigma=0.15$，$t=1$. 将区间[0,1]分成 4 个子区间，并从如下分布中选择 4 个随机数：

$$x \sim \mathcal{N}(0,0.25)$$

(a) 利用这些随机数构造[0,1]上的 W_t，S_t. 将 W_t 和 S_t 的图形画出来．(你将得到与真实路径相近的分段线性路径．)

(b) 将区间[0,1]分成 8 个子区间，重复上述过程．

(c) 对于"很小的"$\Delta>0$ 而言，

$$\log\left(\frac{S_t}{S_{t-\Delta}}\right)$$

服从怎样的分布？

(d) 令 $\Delta=0.25$.

$$\frac{\log S_t - \log S_{t-0.25}}{0.25}$$

代表的是什么？它度量的是哪一部分？该随机变量如何随时间发生变化？

(e) 令 $\Delta=0.000\,001$. 下面的随机变量将如何随时间发生变化？

$$\frac{\log S_t - \log S_{t-\Delta}}{\Delta}$$

(f) 如果 $\Delta \to 0$，下面的“随机变量”的路径会发生什么变化？

$$\frac{\log S_t - \log S_{t-\Delta}}{\Delta}$$

(g) 你认为上一个问题中的变量是一个随机变量吗？

2. 证明，当$\{Y(t), t\geqslant 0\}$满足

$$Y(t) = \exp\left\{cW(t) - \frac{c^2 t}{2}\right\} \tag{7.66}$$

时，它是鞅，其中 c 是任意常数．$E[Y(t)]$等于多少？

121~122

3. 考虑等式 $X_t = X_0 + \mu \mathrm{d}t + \sigma W_t$，$Y_t = \mathrm{e}^{X_t}$. 假设 $\mu=0.1$，$\sigma=0.15$，时间 $t=0.5$. 利用随机正态变量W_t得出X_t和Y_t的路径．计算X_t和Y_t的数学期望．画出X_t和Y_t的直方图，并判断它们服从什么分布．

第8章 维纳过程、列维过程及金融市场上的罕见事件

8.1 引言

在普通交易日的每一个瞬间，都有3种可能的状态：价格上升、下降或不变．事实上，流动的金融工具的价格变化很少超过最低合约报价．所以，连续时间的3种状态金融资产定价方法非常贴合实际，前提是不考虑那些罕见事件．但是，在大多数资本市场和衍生品市场中，都会时不时出现极端行为，这时候我们特别需要准确的定价方法．

怎样的事件才算极端或罕见呢？金融市场中大幅动荡等同于“罕见”吗？本章我们将阐明罕见事件的概率结构，并同维纳过程的行为进行对比．我们还将讨论维纳过程能够描述哪类事件，从而引出罕见事件的特征．

罕见事件与所观测价格过程的间断性有关．这和大幅动荡不同．在连续时间的随机过程中，方差或波动率增大，都会导致波动幅度变大．

罕见事件的波动幅度、发生概率关于观察时间间隔的大小有特别的变化规律，与普通事件不同．具体而言，观察的区间h越小，普通事件的规模也越小，因而它们是“普通”的．在一个月中，可能会发生若干次大幅度的价格变化．在一周中，次数会相应减少．若只观察几分钟，价格大幅跳跃的可能性更是微乎其微．通常，在一分钟普普通通的时间里，没必要担心这类事件发生，这就是“普通”事件的主要特征，当$h\to 0$时，这类事件变得不重要．

另一方面，因为它们是常规平凡的，即便在一个非常小的时间段h，它们的发生概率非零．在很短的时间段内，出现某些“不值得注意”的消息概率总是非零的．

罕见事件则不同．根据定义，它应该不是频繁发生，在连续时间下，这意味着当$h\to 0$时，其发生的概率也趋于0．但是，其事件规模并不减小．1987年那样的市场暴跌是“罕见”的．任意给定一天，挑选一个很短的时间段，观察到这种暴跌的概率是可忽略的，但如果它发生，无论10分钟观察一次还是一整天才观察一次，波动幅度都差不多．

上一章有一个重要结论．在适当的条件下，资产价格不确定部分$\sigma_t \Delta W_t$在小区间h上的方差为

$$E[\sigma_t \Delta W_t]^2 = \sigma_t^2 h \tag{8.1}$$

也就是说，这意味着资产价格变化的不可预测部分期望幅度为$\sigma_t \sqrt{h}$．[⊖]

但要记住标准差是如何获得的：将可能的幅度与相应的概率相乘．得到的是概率乘以

⊖ “期望幅度”仅指变化的绝对值．因为根据定义，我们不知道不可预测部分的方向是怎样的．

事件规模的积．与 h 成正比的方差有两种可能的获得途径，一是概率与 h 有关，而规模是独立的，或者规模与 h 有关，而概率是独立的．[⊖]

第一种情况对应的就是罕见事件，第二种对应的是普通事件．

讨论的相关内容

本章着重讨论罕见事件与普通事件的区别．读者很容易想到，从学术的角度来看这些区别很重要——特别是罕见事件可能会导致资产价格路径不连续．但罕见事件是否有实际应用价值呢？罕见事件的存在是否会影响金融资产定价呢？

这些问题的回答是肯定的．如果资产价格有跳跃，我们必须使用不同的公式，这确实会影响资产定价过程和结果．

124

举个例子，考虑风险管理中的资本需求问题，金融机构应该储备多少资金，以弥补市场中意外变化造成的损失呢？

这取决于有多少价值的资产是有风险的，有多种度量风险值的方法，它们都会关注当其资产价值出现极端变化时，投资组合的价值改变了多少．

此时一个重要的任务是确认有没有罕见事件可能使价格发生跳跃，如果没有，风险值能够使用正态分布进行计算，价格变化可用正态分布的随机变量来模拟，且在适当条件下，风险值也是正态分布的，然后就能直接得到某些极端价格运动后的损失及相应概率．

另一方面，如果这种偶然跳跃是资产价格变化的系统自带成分，那么风险值的计算将变得更为复杂，要给极端情况的损失值赋予相应的概率，我们还要对“罕见”事件建模．

8.2 两个初始模型

连续时间的资产定价模型有两块基石：一是维纳过程，即布朗运动，它是路径连续的随机过程．如果市场上主要是“普通”事件，“极端”波动仅偶然发生且符合正态分布尾部对应的概率，那么就适合用维纳过程．第二是泊松过程，可以对罕见事件导致的系统性跳跃建模，泊松过程是路径非连续的过程．

将以上两种过程用适当的方式结合起来，我们就能得到满足相关需求的模型．

在讨论罕见和普通事件前，本节先回顾一下这两部分基础内容．

8.2.1 维纳过程

在连续时间下，“普通”事件可以用维纳过程（即布朗运动）进行建模．假如随机过程 W_t 的路径只连续变化，用维纳过程是合适的．对于维纳过程，在小区间 h 上，通常只会看到 W_t 变化一个较小的值，这与事件的普通性相一致．

有几种不同的方法可以讨论维纳过程．

⊖ 或者两者都与 h 有关．

一种方法是之前介绍过的，随机变量ΔW_{t_i}在每个时刻

$$0 = t_0 < t_1 < \cdots < t_n = T \tag{8.2}$$

只能取$\sqrt{h}$或$-\sqrt{h}$两个可能的值，其中对所有i，有

$$t_i - t_{i-1} = h \tag{8.3}$$

假定当$i \neq j$时，ΔW_{t_i}和ΔW_{t_j}相互独立，那么求和项

$$W_{t_n} = \sum_{i=1}^{n} \Delta W_{t_i} \tag{8.4}$$

当n趋于无穷大时将弱收敛于维纳过程，这意味着维纳过程是右端求和项一种有效的近似方法．㊀

在该定义中，维纳过程是通过对独立同分布的随机变量求和再求极限得到的．注意，这些增量的可能结果都是时间间隔h的函数，当$h \to 0$时，W_t每次的变化量减小．由这种方法，我们也看到维纳过程在每个时刻是高斯(正态)分布的．

125

我们还可以通过连续平方可积鞅来定义维纳过程．假定W_t是一个有有限方差的连续过程，㊁给定信息集族$\{I_t\}$，增量不可预测．㊂然后根据列维的一个著名定理，这些性质足以保证W_t的增量满足正态分布，且均值为0，方差为$\sigma^2 \mathrm{d}t$.

维纳过程的鞅方法的正式定义如下：

定义 18 相对于信息集族$\{I_t\}$，如果随机过程W_t满足下列条件：

1) W_t是平方可积鞅，$W_0=0$且

$$E[(W_t - W_s)^2] = t - s, \quad s \leqslant t \tag{8.5}$$

2) W_t的轨迹在任意时间t连续．

那么称W_t是维纳过程．

该定义表明维纳过程具有以下性质：

- 由于W_t是鞅，其增量互不相关，因为鞅的增量具有不可预测性．
- W_t的均值为0，因为其初始值为0，且每个增量的均值也等于0.
- W_t的方差为t.
- 最后，该过程是连续的，也就是说，在无穷小区间上，W_t的位移是无穷小的．

在以上定义中，并没有直接说增量是正态分布的．在使用鞅方法定义时，正态性包含在定义的假设中．㊃

在资产价格增量不可预知但价格关于时间连续变化时，维纳过程是非常合适的模型工具. 在继续讨论它之前，我们先澄清一个疑虑．

㊀ 随着n趋于无穷大，右端表达式将会是大量相互独立且无穷小的随机变量的和．在某些条件下，该和式的分布将近似于正态分布，这属于中心极限定理的典型应用，也是连续时间下的弱收敛的应用．

㊁ 即平方可积．

㊂ 这也意味着增量彼此不相关．

㊃ 这就是著名的列维定理．

维纳过程还是布朗运动?

读者可能已经注意到我们将随机过程 W_t 称为布朗运动. 布朗运动和维纳过程到底指的是同一个概念，还是有区别呢?

之前我们用平方可积鞅 W_t 定义了维纳过程. 但并没有说 W_t 服从何种分布.

现在我们来定义布朗运动.

定义 19 若随机过程 B_t，$t\in[0,T]$，满足下列条件

- 该过程的初始点为 0，$B_0=0$.
- B_t 有形式平稳、独立的增量.
- B_t 关于 t 连续.
- 增量 B_t-B_s 服从正态分布，均值为 0，方差为 $|t-s|$:

$$(B_t-B_s)\sim\mathcal{N}(0,|t-s|) \tag{8.6}$$

则称 B_t，$t\in[0,T]$是布朗运动.

该定义和维纳过程有许多相似之处，但有一个关键的区别：W_t 是鞅，而 B_t 没有类似要求，而且，可以推断 B_t 是服从正态分布的.

这些看上去是非常重要的区别，读者因此可能认为 W_t 比布朗运动更一般，因为没有关于其分布的设定.

126

这一感觉是错误的，著名的列维定理告诉我们这两个过程没有任何区别.

定理 4 关于信息族 I_t 的维纳过程是布朗运动.

该定理非常简洁，维纳过程和布朗运动这两个词可以相互替换，所以在后面的章节中这两个概念没有差别.

8.2.2 泊松过程

现考虑一种很不一样的随机环境. 假定 N_t 表示一个金融市场中 t 时刻以前的极端震荡总次数，这些重大事件都是不可预知的.

N_t 的增量只有两种可能的取值，等于 0 表示没有新的重大事件发生，等于 1 表示发生了一个重大事件. 已知重大事件是“罕见的”，等于 1 的增量很少出现.

我们用符号 $\mathrm{d}N_t$ 表示 N_t 在长度为 $\mathrm{d}t$ 的无穷小时间间隔内的变化量. $\mathrm{d}N_t$ 的取值和对应概率如下:[⊖]

$$\mathrm{d}N_t=\begin{cases}1 & \text{概率为 } \lambda\,\mathrm{d}t\\ 0 & \text{概率为 } 1-\lambda\mathrm{d}t\end{cases} \tag{8.7}$$

N_t 的增量在无穷小区间 $\mathrm{d}t$ 有两个可能的取值. 这里与布朗运动有一个关键区别，即泊松过程的增量大小与 $\mathrm{d}t$ 无关. 但是，其增量各取值的概率是 $\mathrm{d}t$ 的函数. 随着观察间隔趋

⊖ 这里使用$\mathrm{d}N_t$和 $\mathrm{d}t$ 而不是ΔN_t和 h 是具有象征意义的. 在后续章节中，符号$\mathrm{d}N_t$和 $\mathrm{d}t$ 的含义将会更加清楚.

于 0，布朗运动的增量越来越小，[1]而 N_t 的变化量大小不变.

读者可能会想起来了，N_t 是泊松计数过程. 假定这些事件在 dt 发生的概率是 λ，定义

$$M_t = N_t - \lambda t \tag{8.8}$$

则该过程是非连续的平方可积鞅.[2]有趣的是

$$E[M_t] = 0 \tag{8.9}$$

以及

$$E[M_t]^2 = \lambda t \tag{8.10}$$

因此，尽管 M_t 的路径不连续，M_t 和 W_t 的一阶矩和二阶矩有着相同的特征. 具体而言，在长度为 h 的时间区间上，这两个过程增量的方差都与 h 成正比.[3]

我们要强调以下几点.

第一，这两种过程的路径差别非常大. 一个是连续的，另一个是纯跳跃式的.

第二，在很小的区间上，M_t 发生跳跃的概率接近 0. 也就是说，这意味着 M_t 的路径比 W_t 更有规律. 因为泊松计数过程在大多数时候都是保持原数值. 尽管 M_t 有离散的跳跃，但它的变差有界. 另一方面，尽管 W_t 的变化值无穷小，但变化的次数是无穷多的. 因此，它的变差是无界的. 所以，定义积分

127

$$\int_{t_0}^{t} f(W_t)\mathrm{d}W_t \tag{8.13}$$

比定义积分

$$\int_{t_0}^{T} f(M_t)\mathrm{d}M_t \tag{8.14}$$

更加困难.

事实也印证了这一点，黎曼-斯蒂尔切斯定义适用于后一种积分.

8.2.3 例子

图 8-1 展示了一个计算机模拟产生的泊松过程. 首先，选取 $\lambda=13.4$，然后固定 $h=0.001$，再用电脑生成泊松计数过程的轨迹 N_t，$t\in[0,1]$. 我们可以总结出泊松分布路径的以下特点：

- 轨迹总体向上倾斜.（所以，N_t 不是鞅.）
- 发生变化时，向上跳跃 1 个单位.
- 两次跳跃之间轨迹为常数.

---

[1] 速度正比于$\sqrt{h}$.

[2] M_t被称为补偿泊松过程. λt 被称为补偿项，它补偿了N_t的正向趋势，并将其变为“无趋势”过程M_t.

[3] $\mathrm{d}M_t$方差的计算方法如下：

$$E[\mathrm{d}M_t]^2 = 1^2\lambda\mathrm{d}t + 0^2[1-\lambda\mathrm{d}t] \tag{8.11}$$

因此有

$$E[\mathrm{d}M_t]^2 = \lambda\mathrm{d}t \tag{8.12}$$

这一点很有启发性，因为我们不知道是否能把 $\mathrm{d}M_t$这类增量看做类似于标准随机变量的量. 为进行准确讨论，我们需要从时间区间的有限细分出发，呈现一些极限讨论.

• 在该特例中，有 12 次跳跃，与均值 λ 很接近．

图 8-2 展示了由泊松过程和维纳过程得到的混合过程的路径．先由泊松过程生成轨迹，然后利用电脑生成具有方差 $h=0.001$ 的标准维纳过程，最后把两个轨迹结合起来．

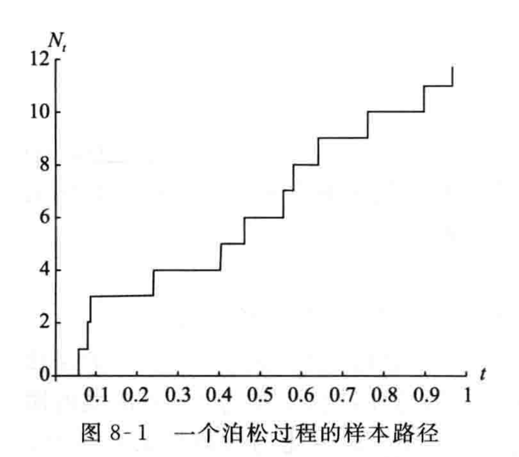


图 8-1　一个泊松过程的样本路径

图 8-2　一个泊松过程与维纳过程的混合过程的样本路径

从该轨迹中我们可以总结出以下特点：

• 该路径偶尔发生跳跃，这来自于泊松过程．

• 在两次跳跃之间，该过程不是常数，它随机波动，这来自于维纳过程．

• 维纳过程产生的噪声明显小于泊松过程产生的跳跃．如果我们选择的维纳过程方差比较大，这一点可能发生变化，那样的话，就很难区别是泊松过程产生的跳跃还是维纳过程产生的波动了．

8.2.4　列维过程

列维过程是所有具有平稳、独立增量的随机过程的统称．列维一辛钦定理根据所包含的过程的特征总结出了列维过程的特征．它告诉我们，存在测度 ν，对所有 $u\in\mathbf{R}$ 和 t 非负，列维过程的特征函数可写为

$$E(\mathrm{e}^{iuX_t}) = \exp(t\phi(u)) \tag{8.15}$$

其中

$$\phi(u) = i\gamma u - \frac{1}{2}\sigma^2 u^2 + \int_{-\infty}^{+\infty} (\mathrm{e}^{iuy} - 1 - iuy1_{\{|y|\leqslant 1\}})\mathrm{d}\upsilon(y) \tag{8.16}$$

这里 γ 和 σ 是实数，ν 是 $\mathbf{R}$ 上的测度，满足 $\nu(0)=0$，且

$$\int_{-\infty}^{+\infty} \min(1,x^2)\mathrm{d}\upsilon(x) \tag{8.17}$$

有界．假定列维过程 $\{X_t\}_{t\geqslant 0}$ 为以下形式：

$$X_t = (r-q+\omega)t + Z_t \tag{8.18}$$

128　该过程包含 ω 控制的漂移项和纯跳跃成分 $\{Z_t\}_{t\geqslant 0}$．在方差伽马过程中，纯跳跃部分的列维测度可以写作 $\mathrm{d}\upsilon(y)=k(y)\mathrm{d}y$，其中 $k(y)$ 为

$$k(y)=\frac{e^{-\lambda_p y}}{\nu y}1_{y>0}+\frac{e^{-\lambda_n |y|}}{\nu |y|}1_{y<0} \tag{8.19}$$

且

$$\lambda_p=\left(\frac{\theta^2}{\sigma^4}+\frac{2}{\sigma^2 \upsilon}\right)^{\frac{1}{2}}-\frac{\theta}{\sigma^2}$$

$$\lambda_n=\left(\frac{\theta^2}{\sigma^4}+\frac{2}{\sigma^2 \upsilon}\right)^{\frac{1}{2}}+\frac{\theta}{\sigma^2}$$

8.2.5 回到罕见事件

相比于常规发生的事件，从定义上看，罕见事件意味着发生变化的规模更大，这种分类方法看起来很明显，但其实并不容易区分．对于一个由维纳过程驱动的随机微分方程，其中的维纳过程在长度为 h 的小区间上的累积反映了长度为 h 的小区间上方差为 $\sigma^2 h$ 的无法预料的价格变化，其中 σ 可能由已知信息得出，此外，这些无法预料的价格变化服从正态分布．

正态分布的尾部延伸到无穷，与此对应，当 h 很小但非零时，发生大幅度价格变化的概率是大于零的，所以，当 h 非零时，维纳过程好像也适合在随机微分方程中对"大"事件建模．那么，为什么我们还要对"罕见"事件或者大事件进行额外的讨论呢?

用维纳过程描述罕见事件存在以下问题：随着 h 趋于 0，正态分布的尾部变得越来越薄，在极限 $h=0$ 时，尾部完全消失，此时整个分布集中于零点．这是意料之中的，因为维纳过程以概率 1 连续，随着 $h\to 0$，用维纳过程表示的意外价格变化的规模将变得越来越小．如此看来，用维纳过程对那些小区间内发生巨大变化的情况进行建模是不合适的． 129

我们需要添加一个干扰项，使得价格能在极小的区间里发生大规模变化．也就是说，我们需要一个可能跳跃的随机过程，且这样的过程具有与 h 无关的变化值，随着 h 减小，其变化值的大小不变．

所以，"罕见"事件应用于例子中偶发的跳跃．

许多衍生品市场都存在价格跳跃，在现货市场这种现象更常见，任何一条新闻都可能影响某标的货物的价格，例如，关于玉米的报道可能使相应的期货价格发生跳跃．但在金融衍生品市场中，新闻影响的效果相对不明显，单独一条新闻对价格或利率或外汇衍生品的影响要小得多，虽然存在影响．

接下来几节我们将描述普通和罕见事件的特征，并学习如何对有偶尔跳跃的价格序列进行建模．

8.3 离散时间上的随机微分方程

要进一步分析普通和罕见事件，最好的办法是研究有限区间上的随机微分方程．[⊖] 130

⊖ 在第 7 章中，我们使用了一些近似来获得离散区间上的 SDE. 对于很小但非无穷小的 h，这种方程只在近似意义下成立．

仍考虑第 7 章中观察间隔为 h 的随机微分方程：

$$S_k - S_{k-1} = a(S_{k-1},k)h + \sigma(S_{k-1},k)\Delta W_k, \quad k = 1,2,\cdots n \tag{8.20}$$

其中 $a(S_{k-1},k)h$ 是漂移项，它表示在下一个区间增量 $S_k - S_{k-1}$ 的期望值是多少 . $\sigma(S_{k-1},k)\Delta W_k$ 是干扰项，决定资产价格中随机成分的值 . 我们已证明在适当假设下，干扰项 $\sigma(S_{k-1},k)\Delta W_k$ 的方差与区间长度 h 成正比 . $\sigma(S_{k-1},k)^2$ 是比例系数 .

为了更具体地认识普通事件和罕见事件，我们还需增加一个假设 .㊀

假定 4 ΔW_k 的可能取值为有限个，将可能的取值和相应的概率记为㊁

$$\sigma_k \Delta W_k = \begin{cases} \omega_1 & 概率为\ p_1 \\ \omega_2 & 概率为\ p_2 \\ & \cdots \\ \omega_m & 概率为\ p_m \end{cases} \tag{8.21}$$

虽然将发生哪一事件是未知的，但所有人员都知道可能发生的事件集 . 一个 ω_i 表示干扰项 $\sigma_k \Delta W_k$ 的一个可能结果，而 p_i 代表相应的概率，参数 m 是可能结果的总数，为整数 .㊂

有两种类型的 ω_i，前 3 个表示“普通”的结果 . 例如，ω_1 表示一次价格上升，ω_2 表示一次下降，ω_3 代表没有变化 . 在现实中，这些显然是资产价格的常规变化 .

其余的 ω_4，ω_5，…，则表示各种罕见的特殊事件 . 例如，假设标的证券是一种关于谷物期货的衍生品，ω_4 可能表示一次重大干旱的影响，ω_5 可能表示一次玉米特大丰收的影响，等等 . 显然，如果这些可能性代表极端价格变化，它们是稀有的，且必然导致比普通波动规模更大的价格变化 . 除了这些罕见事件外，价格变化都是普通事件 ω_1、ω_2、ω_3 的结果.

这一设置将在下一节中用于确定罕见事件的概率结构 .

8.4 罕见事件和普通事件的特征

在上一章的假定 1～3 下，我们证明了一个重要结论：$\sigma_k \Delta W_k$ 的方差为

$$E[\sigma_k \Delta W_k]^2 = \sigma_k^2 h \tag{8.22}$$

该方差与观察间隔 h 成比例，其中在给定信息集 I_{k-1} 情况下 σ_k 是一个已知参数 .

结合假定 4，我们能够进一步研究该结论，通过上述方法我们能够得到普通和罕见事件简洁的特征描述，不过你也许会对繁多的记号感到厌烦 . 但是，这个结论很实用，这一点代价是值得的 .

131

根据假定 4，ΔW_k 只有有限个可能的取值，记为 ω_i，且 ω_i 对应的概率为 p_i，我们可

㊀ 这里遵循 Merton(1990).

㊁ 我们引入这一假设有两个原因：首先，稀有事件和普通事件的差别在可能性有限时更好介绍 . 其次，金融市场中资产定价通常会产生二叉树或三叉树 . 产生二叉树时，市场参与者假设价格在任意瞬间都有两种可能变化 . 产生三叉树时，可能变化增加到 3 种 . 因此在实际操作中，可能情形的总数无论如何都是有限的 .

㊂ ω_i 和 p_i 都可以认为是取决于信息集 I_k 的 . 如果这样假设，就必须加上下标 k，而这会使公式过于冗长 . 为避免这种麻烦，我们令 ω_i 和 p_i 与 k 独立 .

将方差写为

$$V[\sigma_k \Delta W_k] = \sum_{i=1}^{m} p_i \omega_i^2$$

根据上一章的重要命题，这意味着

$$\sum_{i=1}^{m} p_i \omega_i^2 = \sigma_k^2 h \tag{8.23}$$

其中参数 m 是可能状态的数量．等式(8.23)左端是方差的加权平均，其中各状态对应的干扰项的均值是0，权重是各状态对应的概率．[㊀]

现在，(8.23)左端是 m 个有限非负项的和．如果求和结果与 h 成正比，且每项都为正或0，那么每一项都与 h 成比例或等于0. 也就是说，可将 $p_i\omega_i^2$ 写作

$$p_i \omega_i^2 = c_i h \tag{8.24}$$

其中 $c_i > 0$ 是比例系数．[㊁]

(8.24)式表明所有项关于 h 都是线性函数．然后我们可以把 p_i 和 ω_i 看成两个 h 的函数，而它们的乘积与 h 成比例．即

$$p_i = p_i(h) \tag{8.25}$$

和

$$\omega_i = \omega_i(h) \tag{8.26}$$

满足

$$p_i(h)\omega_i(h)^2 = c_i h \tag{8.27}$$

参考 Merton(1997)，将 $p_i(h)$ 和 $\omega_i(h)$ 设为指数形式的函数，

$$\omega_i(h) = \bar{\omega}_i h^{r_i} \tag{8.28}$$

以及

$$p_i(h) = \bar{p}_i h^{q_i} \tag{8.29}$$

其中 r_i 和 q_i 是非负常数，$\bar{\omega}_i$ 和 $\bar{p}_i$ 是可能与 i 和 k 有关的常数，但与观测区间长度 h 无关．

图8-3给出了一些 h^{r_i} 的图形．有3个例子：$r_i=1$ 的情况(对这种特殊情况不作讨论)，$r_i=0.5$ 的情况以及 $r_i=\frac{1}{3}$ 的情况．可以看到，当 h 较小时，$h^{r_i}>h$.

根据(8.28)式和(8.29)式，事件的规模和概率都可能与区间长度 h 有关．当 r_i

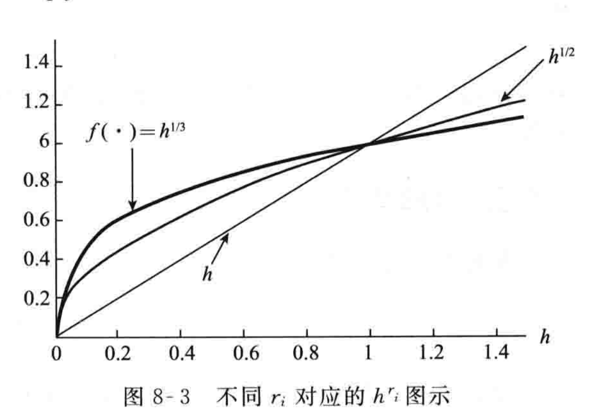


图8-3　不同 r_i 对应的 h^{r_i} 图示

㊀ 我们通过添加下标 k 证明 ω_i、p_i 与信息的潜在相关性．

㊁ 通常 c_i 也是依赖于 k 的．为保持记法的简洁，我们略去了下标 k.

和 q_i 非 0 时，随着 h 增大，价格变化的规模和概率都增大．

我们使用参数 r_i 和 q_i 来描述普通事件和罕见事件的特征，这些参数均是非负的．当观察间隔变小时，r_i 决定了事件规模收敛于零的速度，q_i 决定了概率收敛于零的速度．当然 r_i 或 q_i 有可能为 0，但是它们不能同时为 0.[⊖]

132

我们现在将具体介绍对于普通和罕见事件，r_i 和 q_i 有哪些区别．

(8.24)中 ΔW_k 的方差由以下项组成：

$$p_i\omega_i^2 = \bar{\omega}_i^2\bar{p}_i h^{2r_i}h^{q_i} \tag{8.30}$$

同时我们知道每个 $p_i\omega_i^2$ 都与 h 成比例．

$$p_i\omega_i^2 = c_i h \tag{8.31}$$

因此

$$\bar{\omega}_i^2\bar{p}_i h^{2r_i+q_i} = c_i h \tag{8.32}$$

但这意味着

$$q_i + 2r_i = 1 \tag{8.33}$$

和

$$c_i = \bar{\omega}_i^2\bar{p}_i \tag{8.34}$$

所以参数 q_i、r_i 必须满足限制条件

$$0 \leqslant r_i \leqslant \frac{1}{2} \tag{8.35}$$

和

$$0 \leqslant q_i \leqslant 1 \tag{8.36}$$

我们发现两有种特殊的情况，即

$$r_i = \frac{1}{2}, \quad q_i = 0 \tag{8.37}$$

和

$$r_i = 0, \quad q_i = 1 \tag{8.38}$$

第一种情况对应的就是我们称为“普通”的事件，第二种对应的就是我们称为“罕见”的事件．

8.4.1 普通事件

普通事件的条件是

$$\frac{1}{2} \geqslant r_i \geqslant 0 \tag{8.39}$$

为了理解其含义，考虑当 $r_i=\frac{1}{2}$时会发生什么．

⊖ ω_i和p_i的乘积必须与 h 成正比．如果r_i和q_i同时为 0，其乘积将与 h 无关，这是不允许的．

首先，我们注意到 q_i 等于 0.⊖ ω_i 可写为

$$\omega_i = \bar{\omega}_i h^{1/2} = \bar{\omega}_i \sqrt{h} \tag{8.41}$$

由此可以发现，当 $r_i = 0.5$ 时，随着观察间隔 h 减小，事件的规模变小．另一方面，它们的概率不依赖于 h. 虽然变化很小，但随着 h 缩小，它们的概率恒为常数，它们是“普通的”.

现假定 ΔW_k 所有可能的结果都属于这种类型且 $r_i = 0.5$，那么过程 W_t 的轨迹将有一些有趣的性质．

1. 连续路径

如果不存在罕见事件，那么对于所有的 ω_i 均有 $r_i = 0.5$，它们的规模

$$\omega_i = \bar{\omega}_i \sqrt{h} \tag{8.42}$$

随着 h 减小而减小．同时，当 h 趋于 0 时，ω_i 的值相互靠近，这意味着过程 W_k 在极限情况下是连续的，ΔW_k 移动的距离接近于 0.

$$\lim_{h \to 0} \omega_i = \lim_{h \to 0} \bar{\omega}_i \sqrt{h} \tag{8.43}$$

对每个“普通”事件 ω_i，结论都是如此，我们不用抬手就能画出 W_t 在极限意义下的轨迹，每个增量值都是无穷小． 133

另一方面，由于普通事件的 $q_i = 0$，这些 ω_i 的概率不会随着 $h \to 0$ 而趋于 0，而是与 h 相互独立：

$$p_i = \bar{p}_i \tag{8.44}$$

正是如此，普通事件才能够在连续时间下产生路径．

2. 样本路径的光滑性

对于 $r_i = 0.5$ 时的干扰项，其路径是连续的，但不是光滑的．

首先我们要明白光滑性在确定性函数中的含义：直观地说，如果一个函数不会突变，那它就是光滑的；具体而言，假定我们选取点 x_0，假如对于较小的 h，当 h 趋于 0 时，比率

$$\frac{f(x_0 + h) - f(x_0)}{h} \tag{8.45}$$

始终是有限的，那么 $f(x)$ 在 x_0 点是光滑的．也就是说，如果函数在某点有导数，那么函数在该点是光滑的．

对于 W_t 等非确定性函数，光滑的意义是否相同呢？

在我们讨论的特例中，ΔW_k 的可能取值有 m 个，m 为有限数．这些事件的规模均与 $h^{\frac{1}{2}}$ 成正比．也就是说，随着时间推移，新事件对应的价格变化是 $\sqrt{h}$ 阶的．

⊖ 因为有

$$q_i + 2r_i = 1 \tag{8.40}$$

且 q_i 非负．

在任何时刻，对某个 i，意外价格变化的速度可以写为

$$\frac{f(x_0+h)-f(x_0)}{h} \tag{8.46}$$

取极限

$$\lim_{h\to 0}\frac{W_{t+h}-W_t}{h}=\lim_{h\to 0}\frac{\omega_i}{h} \tag{8.47}$$

将 ω_i 替换，

$$=\lim_{h\to 0}\bar{\omega}_i\frac{h^{\frac{1}{2}}}{h} \tag{8.48}$$

$$=\bar{\omega}_i\lim_{h\to 0}\frac{1}{h^{\frac{1}{2}}}\to\infty \tag{8.49}$$

这意味着随着 h 减小，W_t 变化的速度趋于无穷大，资产价格会变得连续但反复无常．（这里我们假设 $\bar{\omega}_i$ 都是正值．）

我们已得到了普通事件路径的结论，下面将讨论罕见事件的路径．

8.4.2 罕见事件

假定对于某一事件 ω_i，参数 r_i 为 0，那么 q_i 为 1. 根据定义，概率为

$$p_i=\bar{p}_i h \tag{8.50}$$

$r_i=0$，$q_i=1$ 的事件 ω_i 是罕见事件，因为根据该等式，随着 $h\to 0$，概率也趋于 0.

另一方面，这类事件的规模可表示为

$$\omega_i=\bar{\omega}_i \tag{8.51}$$

也就是说，规模与间隔长度 h 的大小无关．

下面我们对罕见事件作进一步的讨论．

1. 样本路径

罕见事件扰动项的路径是非连续的．事实上，这些 ω_i 对应的 $q_i=1$，其波动幅度与 h 无关. 随着观测间隔 h 趋于 0，ΔW_t 所产生的值不会变小，意外价格变化的幅度与 h 无关，当这样的罕见事件出现时，W_t 就会发生跳跃．

134

另一方面，若 $q_i=1$，出现跳跃的概率与 h 有关. 当 h 减小时，跳跃的概率也减小，所以尽管轨迹含有跳跃，但并不常见．

显然，如果随机变量 ΔW_k 包含跳跃，其样本路径是不连续的，除了维纳过程，我们还需要别的模型来描述这种随机行为．

2. 进一步讨论

r_i 和 q_i 如果取其他值会怎么样呢？也就是说，对于范围

$$0<r_i<\frac{1}{2} \tag{8.52}$$

和

$$0 < q_i < \frac{1}{2} \tag{8.53}$$

如果 r_i 和 q_i 取这些范围内的其他值，W_t 的样本路径是什么样的?

结果是，对于取值范围内的其他 r_i，q_i，样本路径是连续但非光滑的，就像维纳过程一样．

很容易说明这一点，因为 $0<r_i<0.5$，代表波动幅度的 ω_i 是 h 的函数，随着 $h\to 0$，ω_i 也趋于 0. 由此可以看出，这类事件不是罕见事件．

对于这些事件，其对应的概率也会随着 h 趋于 0，所以这类结果不会经常出现，由于它们的规模会随 h 变小，它们不算是罕见事件．

8.5 罕见事件的模型

如果存在罕见事件，我们应该用什么方法对资产价格建模呢?

我们用一个方程来描述资产价格，将能观察到的价格变化分解为两部分，在当时已知信息的基础上，一部分是可预测的，一部分是不可预测的．在长度为 h 的间隔上，有

$$S_k - S_{k-1} = a(S_{k-1},k)h + \sigma(S_{k-1},k)\Delta W_t \tag{8.54}$$

随着 h 减小，观察间隔变得无穷小，即可得到连续的情况：

$$\mathrm{d}S_t = a(S_t,t)\mathrm{d}t + \sigma(S_t,t)\mathrm{d}W_t \tag{8.55}$$

在后面的章节中，我们将进一步学习随机微分方程，并介绍微分 $\mathrm{d}S_t$ 和 $\mathrm{d}W_t$ 的确切含义.

用同一形式的方程来表示有罕见事件的模型．罕见事件的发生是超出预期的，它们的方差也与 h 成正比，与维纳过程唯一的差别在于样本轨迹的连续性．因此可以在简单调整后使用相同的随机微分方程，我们需要的是关于随机且不可预测的误差项 $\mathrm{d}W_t$ 的表示．

对于罕见事件，即使 h 无穷小，事件的规模也并非无穷小，而概率则小到可忽略．因此，新的干扰项应该能够反映资本价格偶尔出现的跳跃．除此以外，该模型还要足够灵活，以捕捉跳跃发生概率的潜在变化．

我们可以做具体区分．首先，将误差项分为两部分：由之前的讨论可知，资产价格变化包含了持续发生的普通事件和偶然发生的罕见事件，我们将前一部分记为 ΔW_k，后一部分记为 ΔN_k. 为了更准确，我们假定每次跳跃的幅度为 1，在任意时刻 $k-1$，有

$$N_k - N_{k-1} = \begin{cases} 1 & \text{概率为 } \lambda h \\ 0 & \text{概率为}(1-\lambda h) \end{cases} \tag{8.56}$$

135

其中 λ 与 $k-1$ 时刻的信息集无关．令

$$\Delta N_k = N_k - N_{k-1} \tag{8.57}$$

这样代表幅度为 1，发生频率为常数 λ 的跳跃．[⊖]

那么 N_k 可以用泊松计数过程来模拟，泊松计数过程有以下性质：

⊖ 用相应的概率 λh 除以 h 就可以得到区间 h 内跳跃发生的频率．

1）在小区间 h 内，至多发生一次事件的概率非常接近于 1.

2）t 时刻的已知信息无法预测下一个 h 时间内是否会发生跳跃事件．[⊖]

3）事件的发生频率为常数 λ.

实际上，只有泊松过程同时满足以上条件，它似乎是模拟不连续跳跃的极好选择．不过我们也许还需要进行两项修正．

第一，对一项特定的资产价格而言，跳跃的发生频率可能是随时间变化的，而泊松过程中发生的频率为常数，不符合，故需要调整．

第二，N_t 的增量有非零的均值，SDE 方法只能处理均值为 0 的干扰项，所以还需要消除 dN_t 的均值．

考虑以下修正变量

$$I_t = (N_t - \lambda_t) \tag{8.58}$$

增量 ΔJ_k 的均值将为 0 且不可预测．此外，如果我们用一个(与时间有关的)常数与 J_k 相乘，如 $\sigma_2(S_{k-1},k)$，跳跃的幅度将会与时间相关．因而 $\sigma_2(S_{k-1},k)\Delta J_k$ 是一个表示资产价格跳跃的备用选项．

也就是说，如果市场上某金融工具可能受到罕见事件的影响，其价格变化可用随机微分方程表示为

$$S_k - S_{k-1} = a(S_{k-1},k)h + \sigma_1(S_{k-1},k)\Delta W_t + \sigma_2(S_{k-1},k)\Delta J_t \tag{8.59}$$

当 h 趋于 0，上式变为

$$dS_t = a(S_t,t)dt + \sigma_1(S_t,t)dW_t + \sigma_2(S_t,t)dJ_t \tag{8.60}$$

该随机微分方程同时包含了“普通”事件和“罕见”事件．

最后注意，在每一个时刻 t，跳跃部分 dJ_t 和维纳部分 dW_t 必须相互独立．随着 h 减小，普通事件的规模必须变小，罕见事件则不变．在这些条件下，两类事件必须相互无关，它们的瞬时相关系数为 0.

8.6 有用的矩

“普通”和“罕见”事件的区别之所以重要，还有一个理由．

实际工作中我们需要用到某些矩．在第 5 章中，我们对矩作了定义，它是与所研究的随机过程有关的数学期望．例如，数学期望 $E[X_t]$是一阶矩，方差

$$V[X_t] = E[X_t - E[X_t]]^2 \tag{8.61}$$

是二阶(中心)矩．高阶(中心)矩可表示为

$$E[X_t - E[X_t]]^k \tag{8.62}$$

136

其中 $k>2$.

正如之前所述，矩能够提供随机过程的信息．例如，方差能衡量价格的波动情况．三

⊖ 当 $h\to 0$ 时，该概率为 1.

阶矩能度量价格变化分布的偏度．四阶矩能够度量尾部分布的厚薄．

本节我们将介绍对于无穷小间隔上的普通事件，只有前两阶矩比较关键，高阶矩并不重要．但是对于罕见事件，所有矩都要列入研究范围．

再次考虑之前的设定，即不可预测的随机成分共有 m 种可能结果，每一种记作 ω_i.

不可预测误差项的前两阶矩可表示为⊖

$$E[\sigma_1 \Delta W_k + \sigma_2 \Delta J_k] = [p_1 \omega_1 + \cdots + p_m \omega_m] = 0 \tag{8.63}$$

$$V[\sigma_1 \Delta W_k + \sigma_2 \Delta J_k] = [p_1 \omega_1^2 + \cdots + p_m \omega_m^2] = 0 \tag{8.64}$$

其中 ΔW_k 与 ΔJ_k 无关．

现考虑当所有事件都是普通事件，规模与 $h^{\frac{1}{2}}$ 成正比时，以上矩的大小，即 $q_i = 0$ 的情况．

一阶矩是 m 个值的加权求和，除非结果是 0，否则它与 $h^{\frac{1}{2}}$ 成正比：

$$E[\sigma_1 \Delta W_k] = h^{1/2}[p_1 \bar{\omega}_1 + \cdots + p_m \bar{\omega}_m] \tag{8.65}$$

若将它除以 h，可以得到不可预料的价格变化的平均速率．显然，当 h 较小时，$\sqrt{h}$ 大于 h，而表达式

$$\frac{E[\Delta W_k]}{h} \tag{8.66}$$

随 h 减小而增大．我们可以得出结论：当一阶矩不等于 0 时，它很大且不可忽略，即使观察间隔 h 非常小．

二阶矩亦是如此，价格变化不可预测部分的方差包含 ω_i^2 项．当 ω_i 属于普通型时，其大小与 $\sqrt{h}$ 成正比．因此，方差与 h 成正比：

$$V[\sigma_1 \Delta W_k] = h\Big[\sum_{i=1}^{m} p_i \bar{\omega}_i^2\Big] \tag{8.67}$$

用它除以 h，可以得到平均方差．可以看到，除以 h 后 h 会被消去，随着 h 趋于 0，平均方差恒为常数．

这意味着随着 $h \to 0$，方差不会变得可以忽略．所以对于普通事件，方差包含了无穷小区间 h 上随机变化的重要信息．

现考虑高阶矩

$$E[\sigma_1 \Delta W_k]^n = [p_1 \omega_1^n + \cdots + p_m \omega_m^n] \tag{8.68}$$

其中 $n > 2$.

这里，我们研究的事件是普通型．对 ω_i 求 n 次方，可得

$$\omega_i^n = \bar{\omega}_1^n (h^{1/2})^n \tag{8.69}$$

当 $n > 2$，h 较小时，有

$$h^{n/2} < h \tag{8.70}$$

⊖ 在本节余下部分，$\sigma_i(S_t, t)$，$i=1$，2 将被简写为 σ_i.

所以，将高阶矩除以 h，相应的比率为

137

$$\frac{E[\sigma_1 \Delta W_k]^n}{h} = h^{(n-2)/2} \sum_{i=1}^{m} \bar{\omega}_i^n \tag{8.71}$$

该比率与 h 正相关．随着 h 趋于 0，$h^{\frac{(n-2)}{2}}$ 也趋于 0.[⊖]

所以，当 h 很小时，对普通事件而言，意外价格变化的高阶矩并不包含什么有用信息．我们只需要一个双参数的概率模型，就足以涵盖价格变化的所有信息．其中一个参数是一阶矩，另一个是方差．如果不存在罕见事件，那么最佳选择是维纳过程．

若存在，情况就不一样了．

假设所有事件都是稀有型．根据定义，罕见事件的 ω_i 大小与 h 无关．其二阶矩为

$$E[\sigma_2 \Delta J_k]^2 = h\Big[\sum_{i=1}^{m} \omega_k^2 \bar{p}_i\Big] \tag{8.72}$$

其中 ω_i 与 h 无关．将右端项除以 h，结果中将不含 h. 所以，方差不可忽略．这一点与维纳过程是相同的．

高阶矩可表示为

$$E[\sigma_2 \Delta J_k]^n = h\Big[\sum_{i=1}^{m} \omega_i^n \bar{p}_i\Big] \tag{8.73}$$

对于罕见事件，其概率与 h 成比例，而 h 能被约掉．当 $n>2$ 时，高阶矩也与 h 成比例，若我们将 ΔJ_k 的高阶矩除以 h，得到的比率不会随着 $h\to 0$ 而减小．与维纳过程不同，ΔJ_k 的高阶矩在观察间隔无穷小时不能忽略．这意味着如果价格受罕见事件影响，高阶矩能向市场参与者提供有用的信息．

以上讨论说明了在什么情况下，将随机微分方程中的干扰项局限为维纳过程是合适的. 如果你有足够信心认为金融市场上的某类波动属于普通型，那么你只需要前两阶矩，就能作出合理的近似．dW_t 的正态性假设与最终结果有小的差异，因为我们只使用了前两阶矩来得到数据. 但是如果罕见事件也是整个系统的一部分，只使用维纳过程就不合适了.

8.7 小结

在后面两章，我们将正式地介绍随机微分方程，本章和上一章为学习这部分内容打下了基础．资产价格多样化的动态变化总能用随机微分方程来表示，

$$dS_t = a(S_t, t)dt + [\sigma_(S_t, t)dW_t + \sigma_2(S_t, t)dJ_t] \tag{8.74}$$

其中右端第一项是 S_t 变化的期望值，括号内是随机部分，根据 t 时刻的已知信息无法预测. 由于随机微分方程尚未正式定义，我们经常用 ΔS_k 和 ΔW_k 表示小的增量．

随机微分方程中不可预测的部分由两项组成，dW_t 表示持续发生但规模很小的事件，dJ_t 表示规模很大但难得一见的事件．

⊖ 当 n 大于 2 时，h 的指数为正．

在小区间内，随机变量 W_t 可以被一阶和二阶矩充分描述．高阶矩不能提供任何额外信息．因此，假设 W_t 是维纳过程为这类事件提供了良好的近似．

138

罕见事件不能用正态分布描述，如果它们可能影响金融市场，不可预测的项应该被 dJ_k 过程所补足．泊松过程就具有这类项的良好性质．

由于市场参与者可以随意选择参数 $\sigma_1(S_t,t)$ 和 $\sigma_2(S_t,t)$，维纳过程和泊松过程的结合体可以表示任意可能影响金融市场的扰动．

8.8 实际应用中的罕见和普通事件

本节我们将介绍在模拟资产价格的实用模型中，普通和罕见事件的区别是如何体现的．它们的区别究竟只是一个理论上的存在，还是会在实际应用中具体体现出来呢？

答案是后者．第 2 章介绍的二叉树定价模型是最好的证明．我们将讨论两种二叉树模型，一种是由“普通”事件的随机项确定的模型，另一种是有“罕见”事件的模型．

首先我们要回顾金融资产价格中的标准二叉树模型．我们研究价格为 S_t 的标的股票．此外，即期汇率 r_t 等过程也可以作为研究对象．

8.8.1 二叉树模型

我们希望将连续时间过程在时间区间 $[0,T](T<\infty)$ 上进行离散化．同时我们希望这种离散化是规则明确且简洁的．和之前一样，我们将长度为 T 的时间区间平均分为 n 个小区间，每段长为 Δ：

$$t_0 = 0 < t_1 < \cdots < t_n = T \tag{8.75}$$

且

$$n\Delta = T \tag{8.76}$$

由此可得离散的时间点 $\{t_i\}$.

下面我们用模型得出 S_t 在这些时间点 t_i 的具体值．为使记号简洁，我们将这些值记作 S_i：

$$S_i = S_{t_i}, \quad i = 0,1,\cdots,n \tag{8.77}$$

在二叉树模型中，一旦达到特定的状态或者节点，也就是每个离散点 i，那 S_i 的下一步变化仅限于上升和下降，且依赖于参数 u_i 和 d_i. ⊖

u_i 和 d_i 如何取值取决于 S_t 将作何种变化．我们将讨论两种情况．

在第一种情况下，u_i 和 d_i 的大小与 Δ 有关，而相应的概率与 Δ 无关，显然，这种情况与“普通”事件对应，且最终可以用维纳过程对应的随机变量来描述．第二种情况相反，u_i 和 d_i 的大小与 Δ 无关，而相应的概率与 Δ 有关，它与“稀有”事件对应，最后会产生泊

⊖ 这些状态被称为“涨”或“跌”，但实际上，两种运动可能是上涨和维持不变或下跌和维持不变．这些项的名字仅具有象征意义．参数 u_i 和 d_i 也可能依赖于该节点或前一节点所观察到的 S_i．这里我们采用更为简单的情形，上涨、下跌的状态相互独立．

松型过程的行为.

8.8.2 普通事件

139 假设瞬时变化率为 μ，瞬时波动率为 σ. 对下面研究的两种情况，我们均假设 S_i 按以下规律变化：

$$S_{i+1}=\begin{cases}u_i S_i & \text{概率为 } p_i \\ d_i S_i & \text{概率为 } 1-p_i\end{cases} \tag{8.78}$$

若 S_i 只受普通事件影响，其变化系数 u_i 和 d_i 可以表示为[⊖]：

$$u_i = e^{\sigma\sqrt{\Delta}} \quad \text{任意 } i \tag{8.79}$$

$$d_i =- e^{\sigma\sqrt{\Delta}} \quad \text{任意 } i \tag{8.80}$$

而概率 p_i 可写为

$$p_i = \frac{1}{2}\left[1+\frac{\mu}{\sigma}\sqrt{\Delta}\right] \tag{8.81}$$

首先，在每个节点 i，u_i、d_i 和 p_i 都取不变的值. 这是因为等式(8.79－8.81)的右端均与 $S_i(i=1, \cdots, n)$无关. 由此可知，离散化的 S_i 在不同时间以相同的规则变化. 显然，这并非必须，我们可以视情况选择更复杂的 u_i、d_i 或 p_i，只要它们和 Δ 的关系与等式(8.79－8.81)中的一样. 在这个特例中，我们可以把 u_i、d_i 的下标 i 去掉.

其次，也是和我们的目的关系更密切的，考虑当 Δ 趋于零时，u_i、d_i、p_i 将如何变化.

从这些参数的定义我们可以看出，随着 $\Delta \to 0$，u_i、d_i 趋于 0. 所以根据(8.78)，当区间无穷小时，S_t 的变化就变得可忽略，但是这些运动的概率趋于常数 $\frac{1}{2}$：

$$\lim_{\Delta\to 0}\frac{1}{2}\left[1+\frac{\mu}{\sigma}\sqrt{\Delta}\right]=\frac{1}{2} \tag{8.82}$$

如果按上述方法设定二叉树模型，其结果与普通事件主导的 S_i 变化毫无矛盾，虽然在小区间上，这种波动持续发生，但规模很小.

8.8.3 罕见事件

现在我们保留二叉树模型的基本框架，仅改变 u_i、d_i、p_i 的定义. 尤其是，我们将改变这些参数与 Δ 的关系.

我们用以下假设代替等式(8.79－8.81)：

$$u_i = \hat{u} \quad \text{任意 } i \tag{8.83}$$

$$d_i = e^{\alpha d} \quad \text{任意 } i \tag{8.84}$$

而上升概率 p_i 为

⊖ 这不是普通事件的唯一特征.

$$p_i = \lambda\Delta \tag{8.85}$$

其中系数 $\lambda(\lambda>0)$和 $\alpha(\alpha>0)$可根据跳跃的期望大小和概率进行选择. $\hat{u}$ $(u\neq1)$也是大于零的常数，它表示 S_i 发生跳跃的情况，d_i 表示没有跳跃的情况.

此类二叉树行为又有哪些特征呢？随着观察间隔 Δ 趋于 0，第一种状态的概率 p_i 趋于 0，另一种状态的概率逼近 1. 这意味着随着间隔减小，价格发生第一种变化的可能性减小. 随着 $\Delta\to0$，S_i 在有限区间内保持一种平稳的路径. 但是即便 Δ 很小，"罕见"事件还是有很小的概率发生，因为根据(8.85)：

$$P(S_{i+1} = \mathrm{e}^{\alpha\Delta}S_i) = 1-\lambda\Delta \tag{8.86}$$

依赖于 Δ，它的概率只是非常接近于 1.

当间隔很小时，有

$$d_i = \mathrm{e}^{\alpha\Delta} \tag{8.87}$$

$$\approx 1 \tag{8.88}$$

其中 Δ 接近于 0.

140

显然，这样定义参数，模型将更符合本章前面介绍的有罕见事件的特征.

8.8.4 累积变化值的特征

上面的讨论是针对离散的双状态过程 S_i，介绍了其变化规模和概率关于 Δ 的可行的建模方法. 我们主要关注了当 Δ 越来越小时，S_i 的下一步变化会怎样.

我们还可以提出一个有趣的问题：不再考虑下一步的变化，随着时间推移，S_i 的累积变化值将有怎样的特征呢？

也就是说，我们不再研究当 i 增加时 S_i 下一步变化的概率，我们更感兴趣的是

$$\frac{S_{i+n\Delta}}{S_i} \tag{8.89}$$

的性质，其中 n 为大于 1 的整数. 即经过 n 次观察之后 S_i 的累积变化值有怎样的特征.

首先，我们要说明研究这种随机变量的理由.

显然，用双状态过程对 S_i 建模对短期的未来是合理的，特别是当 Δ 很小时. 但市场参与者还不知道，当期限较长时，n 个 Δ 之后会发生什么. 例如，专业人士可能想知道到期日的价值 S_T，$t<T$，而不是瞬时价格 S_t，之前的模型并没有介绍这方面内容.

所以，专业人士非常关心到期日价值 S_T 的随机行为，就像关心下一时刻价值一样. 而累积变化值的随机行为与 S_i 的瞬时变化有很大区别，后者只是由 p_i 控制的. 而前者，在经过 n 个观察期之后，S_i 有许多可能的取值，不再是 u_iS_i 或 d_iS_i 两种.

所以，考虑以下比率：

$$\frac{S_{i+n\Delta}}{S_i} \tag{8.90}$$

它与二叉树模型中参数定义方式有关. 我们将用到取值为整数的随机变量 Z，它记录节点 i 到$i+n$ 期间向上运动的次数. 也就是说，从第 i 点开始，如果 S_i 每次都向上运动，那么

$Z=n$. 如果只有一半时间运动向上，那么 $Z=n/2$，以此类推.

我们将研究 S_{i+n}/S_i 取对数后的随机行为，而非该比率本身. 因为这样可以得到 Z 的线性函数.㊀

在进一步研究前，我们将 u_i、d_i、p_i 的下标 i 去掉，因为至少在本节中，它们是常数.

现在我们可以写出：

$$\log \frac{S_{i+n}}{S_i} = Z \log u + (n-Z) \log d \tag{8.91}$$

$$= Z \log \frac{u}{d} + n \log d \tag{8.92}$$

如前面的段落所述，最后我们得到了关于随机变量 Z 的线性函数.

由于是线性函数，我们很容易计算随机变量 $\log(S_{i+n}/S_i)$的均值和方差：

141

$$E\left[\log \frac{S_{i+n}}{S_i}\right] = \log \frac{u}{d} E[Z] + n \log d \tag{8.93}$$

$$V\left[\log \frac{S_{i+n}}{S_i}\right] = \left[\log \frac{u}{d}\right]^2 V[Z] \tag{8.94}$$

而我们知道 $E[Z]$为 np 而 $V[Z]$为 $np(1-p)$.㊁

将其带入：

$$E\left[\log \frac{S_{i+n}}{S_i}\right] = \log \frac{u}{d} np + n \log d \tag{8.96}$$

$$V\left[\log \frac{S_{i+n}}{S_i}\right] = \left[\log \frac{u}{d}\right]^2 np(1-p) \tag{8.97}$$

其中

$$n = \frac{T}{\Delta} \tag{8.98}$$

将 u，d，p 的值带入(8.96)、(8.97)后我们可以得到均值和方差的近似值. 也就是说，将等式(8.79－8.81)中 u，d，p 的值带入，即可得到以下近似值：

$$\log \frac{u}{d} np + n\log d \approx \mu T \tag{8.99}$$

$$\left[\log \frac{u}{d}\right]^2 np(1-p) \approx \sigma^2 T \tag{8.100}$$

这等价于一个$[0,T]$上每步期望大小为 $\mu\Delta$ 的过程，而每一步的波动率等于 $\sigma\sqrt{\Delta}$. 所以这样建模的 S_i，其每步变化的均值和方差都与 Δ 成正比. 这样的随机过程叫做几何过程.

㊀ u_i，d_i是乘积参数. 对数运算将乘积项变为和项，在渐进理论中更好分析. 中心极限定理通常是用和式表示的.

㊁ 期望值很容易计算. 如果有 n 个相互独立的阶段，每个阶段以概率 p 上升，那么总的上升次数期望值为 np. $V[Z]$的计算要复杂一些，单个阶段下 Z 的方差为：

$$p(1-p)^2 + (1-p)(0-p)^2 = p(1-p) \tag{8.95}$$

共有 n 个阶段时，n 个独立运动的总方差为 n 倍的 $p(1-p)$，即 $np(1-p)$.

我们也可以得到 $\log(S_{i+n}/S_i)$ 近似的概率分布．首先，注意 $\log(S_{i+n}/S_i)$ 实际上是原过程取对数后的变化值：

$$\log\left(\frac{S_{i+n}}{S_i}\right)=\log S_{i+n}-\log S_i \tag{8.101}$$

可以证明，如果我们按照(8.79)和(8.80)设定参数，使得模型与普通事件对应，那么当 $\Delta\to 0$ 时，$[\log S_{i+n}-\log S_i]$ 的分布可近似为：

$$[\log S_{i+n}-\log S_i]\sim\mathcal{N}(\mu(n\Delta),\sigma^2\Delta) \tag{8.102}$$

也就是说，$[\log S_{i+n}-\log S_i]$ 近似服从正态分布．

另一方面，如果我们按照(8.83)和(8.84)设定参数，使模型与罕见事件对应，那么当 $\Delta\to 0$ 时，$[\log S_{i+n}-\log S_i]$ 的分布可近似为：

$$[\log S_{i+n}-\log S_i]\sim \text{泊松} \tag{8.103}$$

这是中心极限定理的两个例子．根据该定理，大量随机变量之和往往具有可辨认的近似分布．

为什么在中心极限定理的作用下，这两种事件会导致不一样的结果呢？

原因是，如果想要让随机变量的和收敛到正态分布，求和中的每一项必须都是渐进可忽略的．而渐进可忽略这一条件正是普通事件和罕见事件的差别所在．所以，确定罕见事件的 u_i，d_i，p_i 之后，事件不是渐进可忽略的，而累积变化的极限收敛到泊松分布．

8.9 参考阅读

Merton(1990)中包含了对罕见事件特征的介绍．本章中，我们假设干扰项只有有限个可能取值，该假设极大地简化了讨论，读者如果对本章的理论的进一步讨论感兴趣，可以考虑 Bremaud(1979). Bremaud 采用了鞅方法来讨论点过程的动态变化，这种过程可以看作是泊松过程的一般化． 142

8.10 习题

1. 证明当 $n\to\infty$ 时：

(a) $1\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\to 1$ (b) $\left(1-\frac{\lambda}{n}\right)^n\to e^{-\lambda}$ (c) $\left(1-\frac{\lambda}{n}\right)^k\to 1$

2. 设随机变量 X_n 具有二叉树分布：

$$X_n=\sum_{i=1}^{n}B_i$$

其中 B_i 相互独立，取值为 1(概率为 p)或 0(概率为 $1-p$).

我们可以将 X_n 视为一系列依次发生的随机事件的和，事件为 B_i. 注意到有两个关键的参数，即 p 和 n. 第一项控制各事件的概率，第二项控制事件的数量．问题是，当事件数量趋于无穷时，X_n 的分布是怎样的？有两种有意思的情况，以下是相关问题．

(a) 假设 $n\to\infty$，$p\to 0$，而 $\lambda=np$ 恒为常数．也就是说，$B_i=1$ 的概率随着 n 的增加而趋于 0. 但是，得到 1 的期望频率保持不变．这显然对概率的收敛施加了一定的速度．概率 $P(X_n=k)$ 是多少？写出

关于 p，n，k 的函数公式．

(b) 将 $\lambda=np$ 带人，写出 $P(X_n=k)$关于问题 1 中 3 项参数的函数．

(c) 求证当 $n\to\infty$时得到泊松分布：

$$P(X_n = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

(d) 注意在趋于极限的过程中，p 以某个固定的速度趋于 0，你将如何解释这种极限概率？哪种罕见事件适合于这种概率？

3. 设 T_α 为维纳过程首次到达 α 的时刻，其中 $\alpha>0$，计算 $P(T_\alpha\leqslant t)$.

4. 生成 $n=1000$ 的二叉树随机变量，$p=0.001$. 考虑 B_i，B_i 取值为 1(概率为 p)或 0(概率为 $1-p$)，运行 1000 次，画出 X_i 最终值的柱状图．

143

第9章 随机积分

9.1 引言

在实践应用中，人们对微积分运算感兴趣的原因之一在于需要用它来得到微分方程，而微分方程是用来描述物理现象动态原理的主要工具．一个简单的线性微分方程形如

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = AX_t + By_t, \quad t \geqslant 0 \tag{9.1}$$

其中 $\mathrm{d}X_t/\mathrm{d}t$ 是 X_t 关于 t 的导数，y_t 是外生变量．A 和 B 是参数．㊀

常微分方程是实际建模中的必备工具．例如，工程师可能认为 y_t 与 X_t 的过去值共同决定了 X_t 的未来值，微分方程就可以近似地表达这种关系，因此微分方程有着广泛的应用．㊁

通常，我们通过下列途径获得常微分方程：首先定义导数，显然多数函数，记为 X_t，是有导数的，当导数存在时，用泰勒展开式对 $\mathrm{d}X_t/\mathrm{d}t$ 进行近似，考虑相关理论对应的限制条件，就得到了微分方程．

微积分基本定理表明积分和微分之间有着密切的联系．因为积分表示增量和，微分表示变化速率，当变量 X_t 增加了 $\mathrm{d}X_t$，且初值 $X_0=0$ 时，可以通过下式得到X_t的最新值：

$$\int_0^t \mathrm{d}X_u = X_t \tag{9.2}$$

这表明，任意一个微分方程都有一个对应的积分方程．

然而在随机微积分中是无法使用类似的方法的，因为如果不可预测的消息是持续获得的，且阐述现象动态原理的方程中包含这种噪声，在这种情况下导数将无法定义．

但是在某些条件下，可以获得其积分，这使我们可以用随机微分方程

$$\mathrm{d}X_t = a_t\mathrm{d}t + \sigma_t\mathrm{d}W_t, \quad t \in [0,\infty) \tag{9.3}$$

去替代常微分方程，其中用 $\mathrm{d}X_t$、$\mathrm{d}t$ 和 $\mathrm{d}W_t$ 代替 $\mathrm{d}X_t/\mathrm{d}t$ 表示未来的变化．这些微分是利用新的积分概念定义的．例如，当 h 趋于 0 时，增量

$$X_{t+h} - X_t = \int_t^{t+h} \mathrm{d}X_u \tag{9.4}$$

可以用来定义 $\mathrm{d}X_t$. 我们在前面已经使用过 $\mathrm{d}S_t$、$\mathrm{d}W_t$ 这类微分，但是从未对它们进行过严格的定义，伊藤积分的定义就满足了这一点．

现考虑下列随机微分方程，它表示资产价格 S_t 的动态变化：

$$\mathrm{d}S_t = a(S_t,t)\mathrm{d}t + \sigma(S_t,t)\mathrm{d}W_t, \quad t \in [0,\infty) \tag{9.5}$$

㊀ 如果 $B=0$，就称该方程是齐次的．当 y_t 与 t 相互独立时，该系统是自治系统，否则是非自治系统．

㊁ 例如，工程师可能希望获得 X_t 的某种未来轨迹，则就是在找到合适的$\{y_t\}$后，使 X_t 遵循该轨迹．

对等式两边同时求积分，该等式变为

$$\int_0^t \mathrm{d}S_u = \int_0^t a(S_u,u)\mathrm{d}u + \int_0^t \sigma(S_u,u)\mathrm{d}W_u \tag{9.6}$$

等式右边最后一项是关于维纳过程 W_t 的积分．

目前我们还没有给出(9.6)式右边积分的含义．正如在第 5 到 7 章中讨论的那样，W_t 在小区间 h 内的增量很不稳定，W_t 的平均变化速率等于 $h^{-1/2}$，并随 h 变小而变大，⊖如果这些增量过于不稳定，它们的和会是有限的吗？

本章将解决这些看起来很难的问题．

9.1.1 伊藤积分与随机微分方程

对伊藤积分进行正式定义将会使随机微分方程的概念更加清晰．只要积分

146

$$\int_0^t \sigma(S_u,u)\mathrm{d}W_u \tag{9.7}$$

能够被严格地定义出来，我们就可以对随机微分方程(9.5)中的两端同时求积分得到：

$$S_{t+h} - S_t = \int_t^{t+h} a(S_u,u)\mathrm{d}u + \int_t^{t+h} \sigma(S_u,u)\mathrm{d}W_u \tag{9.8}$$

其中 h 是有限时间区间．

我们可以使用已经在第 7 章与第 8 章中多次使用的有限差分逼近方法，当 h 很小且 $u \in [t,t+h]$时，$a(S_u,u)$ 和 $\sigma(S_u,u)$变化幅度可能不大，尤其是当它们是 S_u 和 u 的光滑函数时. 接下来我们就可以将该等式重新写成：

$$S_{t+h} - S_t \approx a(S_t,t)\int_t^{t+h} \mathrm{d}u + \sigma(S_t,t)\int_t^{t+h} \mathrm{d}W_u \tag{9.9}$$

直接求积分，可以得到有限差分逼近：

$$S_{t+h} - S_t \approx a(S_t,t)h + \sigma(S_t,t)[W_{t+h} - W_t] \tag{9.10}$$

重写成

$$\Delta S_t \approx a(S_t,t)h + \sigma(S_t,t)\Delta W_t \tag{9.11}$$

这就是我们在前几章中经常用到的有限区间内的随机微分方程表达式．将该表达式作为逼近至少有两个理由：首先，$E_t[S_{t+h} - S_t]$ 等于一阶泰勒级数近似：

$$E_t[S_{t+h} - S_t] = a(S_t,t)h$$

其次，当 $u \in [t,t+h]$时，$a(S_u,u)$ 和 $\sigma(S_u,u)$是用 $u=t$ 时刻的值近似的．这两种近似都要求 $a(S_u,u)$ 和 $\sigma(S_u,u)$是光滑的．这表示我们可以得到

$$\mathrm{d}S_t = a(S_t,t)\mathrm{d}t + \sigma(S_t,t)\mathrm{d}W_t \tag{9.12}$$

这意味着，在积分方程

$$\int_t^{t+h} \mathrm{d}S_u = \int_t^{t+h} a(S_u,u)\mathrm{d}u + \int_t^{t+h} \sigma(S_u,u)\mathrm{d}W_u \tag{9.13}$$

⊖ 平均变化速率指 $W_{t+h} - W_t$ 的标准差除以 h. 第 6 章表明，在通常的假设下，不可预测波动的标准差与 $h^{1/2}$ 成正比.

中，右边第二项是在伊藤积分意义下定义的，且 $h\to 0$，

$$\int_t^{t+h}\sigma(S_u,u)\mathrm{d}W_u \approx \sigma(S_t,t)\mathrm{d}W_t \tag{9.14}$$

因此，伊藤积分是无穷小时间区间内对扩散项的近似．

为了使得这些近似有意义，我们必须要首先严格定义关于 W_t 的积分，其次我们要对 $a(S_u,u)$和 $\sigma(S_u,u)$随时间变化的路径添加一些条件，我们不能允许这些 I_t 可测的参数过于不稳定．

9.1.2 实际应用中的伊藤积分

在实践中，伊藤积分的应用少于随机微分方程．从业者几乎从不直接使用伊藤积分去计算衍生品的价格．在后面我们将会提到，衍生品的无套利价格是通过偏微分方程或鞅变换求出的，这两种计算方法都不需要直接利用伊藤积分．

从交易员的角度很难看出这个概念的实际关联．看起来定义伊藤积分只是理论结果而没有实际应用．从业者承认伊藤积分的存在，但他们更倾向于直接使用随机微分方程．

读者们要当心这一点．理解伊藤积分的定义非常重要，原因有二．首先，如前所述，随机微分方程只有在伊藤积分的意义下成立．理解随机微分方程的真实含义需要理解伊藤积分．否则在将随机微分方程应用到实际问题中可能会产生错误．

147

这也给出了伊藤积分相关的第二个原因．由于随机微分方程是定义在无穷小的区间上的，它们在有限区间上的使用可能需要一些近似．事实上，(9.14)中的近似在当 h 不是很小时可能无效．那么就需要利用伊藤积分定义一种新的近似方法．

这一点对应金融衍生品定价非常重要，因为实践中我们都是用有限区间进行计算的．比如“一天”显然不是一个无穷小的区间，随机微分方程在这种区间上的应用可能就需要一些近似．通过伊藤积分得到这些近似的精确形式．

总的来说，从定义在有限区间上的随机微分方程

$$\Delta S_k = a_k h + \sigma_k \Delta W_k,\quad k=1,2,\cdots n \tag{9.15}$$

得出随机微分方程

$$\mathrm{d}S_t = a(S_u,u)\mathrm{d}t + \sigma(S_u,u)\mathrm{d}W_u,\quad t\in[0,\infty] \tag{9.16}$$

等价于利用有意义的 $\int_t^{t+h}\sigma(S_u,u)\mathrm{d}W_u$ 阐释 $\mathrm{d}W_t$，反之亦然．而这只有通过构建随机积分来做到．

9.2 伊藤积分

伊藤积分是定义一段时间内不可数且不可预测的随机增量和一种方法．这种积分无法通过黎曼-斯蒂尔切斯积分的方法获得．

可以看出即使在不久的将来，$\mathrm{d}W_t$ 也是代表一个不可预测的随机变量．维纳过程在 t 时刻的值记为 W_t，是不可数个相互独立的增量和：

$$W_t = \int_0^t \mathrm{d}W_u \tag{9.17}$$

(在 0 时刻维纳过程的值为 0. 因此 $W_0=0$.)这是我们可以写出的最简单的随机积分.

可以对随机微分方程中的干扰项积分得到一个更重要的随机积分：

$$\int_0^t \sigma(S_u, u)\mathrm{d}W_u \tag{9.18}$$

式(9.17)和(9.18)中的积分是不稳定的随机变量之和，当 $\varepsilon>0$ 时，不同时刻的两个变化量 $\mathrm{d}W_t$ 和 $\mathrm{d}W_{t+\varepsilon}$ 也是互不相关的. 那么接下来的问题就是如何定义这类不稳定项的和. 毕竟如此多(不可数)个不稳定项的和很可能是无界的.

再次考虑标准微积分中定义积分的方法.

9.2.1 黎曼-斯蒂尔切斯积分

假设 $F(x_t)$是非随机函数，其中 x_t 是表示时间的确定性变量，$F(\cdot)$连续且可微，其导数为

$$\frac{\mathrm{d}F(x_t)}{\mathrm{d}x_t} = f(x_t) \tag{9.19}$$

当导数 $f(\cdot)$存在时，黎曼-斯蒂尔切斯积分可以被写成两种形式：

$$\int_0^T f(x_t)\mathrm{d}x_t = \int_0^T \mathrm{d}F(x_t) \tag{9.20}$$

等式左边是关于 x_t 的积分，$t\in[0,T]$. 那么这些(不可数个)$f(\cdot)$在任意 x_t 的值乘以无穷小增量 $\mathrm{d}x_t$ 的积可以用于得到积分. 这种记法是黎曼积分的通常记法.

等式右边的记法表示关于 $F(\cdot)$的积分. 利用 $F(\cdot)$的增量获得积分. 我们可以进一步将后面这种记法复杂化. 比如我们会对计算下面的积分感兴趣

148

$$\int_0^T g(x_t)\mathrm{d}F(x_t) \tag{9.21}$$

我们得到的是函数 $g(x_t)$ 关于 $F(\cdot)$的积分.

计算随机变量的期望时也会出现类似的积分. 比如 $F(\cdot)$表示随机变量 x_t 的分布函数，我们可以这样计算 $g(x_t)$当 t 固定时的期望值:[⊖]

$$E[g(x_t)] = \int_{-\infty}^{\infty} g(x_t)\mathrm{d}F(x_t) \tag{9.22}$$

在该积分中，通常 x_t 的变化范围为负无穷到正无穷，利用 $\mathrm{d}F(\cdot)$对 $g(\cdot)$进行平均. 这里 $\mathrm{d}F(\cdot)$表示取对应值的概率.

注意(9.21)和(9.22)之间有一个重要的区别. 在(9.21)中，t 是从 0 变化到 T 的. 在特定时刻 t，x_t 的值是不确定的. 它是随机变量. 这使积分也成为了随机变量.

(9.22)中的积分则不同. t 是常数，积分变量 x_t 从负无穷大到正无穷大. 该积分不是

⊖ 当函数 $g(\cdot)$ 是 x_t 的平方或立方时，该积分就是二阶或三阶矩.

随机变量.

当没有随机变量时，黎曼-斯蒂尔切斯积分是无穷和的极限．只要极限存在，积分就存在．为了强调其与伊藤积分之间的差别，我们再次复习一下黎曼-斯蒂尔切斯方法．

假设我们要计算

$$\int_0^T g(x_t)\mathrm{d}F(x_t)$$

黎曼-斯蒂尔切斯方法计算的基础是用时间点将区间$[0, T]$分割为n个小区间

$$t_0 = 0 < t_1 < \cdots < t_n \tag{9.23}$$

接下来定义有限黎曼和：

$$V_n = \sum_{i=0}^{n-1} g(x_{t_{i+1}})[F(x_{t_{i+1}}) - F(x_{t_i})] \tag{9.24}$$

等式右边是下列项的和：

$$g(x_{t_{i+1}})[F(x_{t_{i+1}}) - F(x_{t_i})] \tag{9.25}$$

其中每一项是$g(x_{t_{i+1}})$与$F(x_{t_{i+1}})-F(x_{t_i})$的积．第一项表示$g(\cdot)$在$x_{t_{i+1}}$处的值．第二项类似增量$\mathrm{d}F(x_t)$．每一项$g(x_{t_{i+1}})[F(x_{t_{i+1}})-F(x_{t_i})]$都可以看做底为$F(x_{t_{i+1}})-F(x_{t_i})$，高为$g(x_{t_{i+1}})$的矩形面积．

V_n是这种矩形面积的和．如果连续的$t_i(i=0, \cdots, n)$彼此距离不远，即$[0,T]$被划分得很好，这就是一种很好的近似．也就是说，如果函数$g(\cdot)$是可积的，那么极限

$$\lim_{\sup_i|t_{i+1}-t_i|\to 0} \sum_{i=0}^{n-1} g(x_{t_{i+1}})[F(x_{t_{i+1}}) - F(x_{t_i}))] = \int_0^T g(x_t)\mathrm{d}F(x_t) \tag{9.26}$$

就是存在的，并且被称为黎曼-斯蒂尔切斯积分．读者可将该等式作为定义．该积分被定义为右边和的极限．[1]和V_n称为黎曼和．[2]

149

9.2.2 随机积分和黎曼和

因此黎曼-斯蒂尔切斯积分可以用底很小而高是变化的矩形近似．对随机积分我们可以用类似的方法吗?

我们可以将随机微分方程写成有限等区间h上的形式进一步讨论：[3]

$$S_k - S_{k-1} = a(S_{k-1},k)h + \sigma(S_{k-1},k)\Delta W_k, \quad k = 1,2,\cdots,n \tag{9.27}$$

假设对(9.27)左边的ΔS_k求和：

$$\sum_{k=1}^{n-1}[S_k - S_{k-1}] = \sum_{k=1}^{n-1}[a(S_{k-1},k)h] + \sum_{k=1}^{n-1}\sigma(S_{k-1},k)[\Delta W_k] \tag{9.28}$$

我们可以使用类似黎曼-斯蒂尔切斯的方法，将随机变量S_t的积分定义为(某种)极限吗?

---

[1] 如果极限收敛．

[2] 有很多矩形可以对曲线进行近似．可以使用同底但高为$g(x_{t_i})$或$g\left(\frac{x_{t_{i+1}}+x_{t_i}}{2}\right)$的矩形．

[3] 考虑等区间可以使$[0,T]$的划分在$n\to\infty$时非常精细．否则要使用$\sup_i|t_i-t_{i-1}|\to 0$.

$$\int_0^T \mathrm{d}S_u = \lim_{n\to\infty}\left\{\sum_{k=1}^{n-1}[a(S_{k-1},k)h]+\sum_{k=1}^{n-1}\sigma(S_{k-1},k)[\Delta W_k]\right\} \tag{9.29}$$

其中 $T=nh$.

(9.28)右边第一项在已知 k 时刻信息时不包含任何随机变量．更重要的是，这是对时间 h 内增量的积分．由定义可知时间是光滑函数且有有限变差．这意味着可以使用黎曼-斯蒂尔切斯方法定义积分[⊖]

$$\int_0^T a(S_u,u)\mathrm{d}u = \lim_{n\to\infty}\sum_{k=1}^{n-1}[a(S_{k-1},k)h] \tag{9.31}$$

但是(9.28)中第二项即使在当已知 I_{k-1} 时，仍包含随机变量．事实上在时刻 $k-1$，项

$$W_k - W_{k-1} \tag{9.32}$$

是随机变量，和

$$\sum_{k=1}^{n-1}\sigma(S_{k-1},k)[W_k - W_{k-1}] \tag{9.33}$$

是关于随机变量的积分．

我们有以下几个问题：

- 应该使用哪种极限？提到这个问题，是因为(9.33)中的和是随机的，其极限形式也应该收敛于某随机变量．黎曼-斯蒂尔切斯方法中使用的确切的极限定义在这里不适用.
- 在什么条件下极限收敛(即(9.33)中的和有有意义的极限)？
- 极限随机变量有什么性质？

150

我们将注意力集中于随机微分方程中误差项决定的特殊积分．可以证明，在某种条件下可以用随机和的均方极限定义随机积分：

$$\sum_{k=1}^{n-1}\sigma(S_{k-1},k)[W_k - W_{k-1}] \tag{9.34}$$

该积分是随机变量．

适用均方收敛意味着和

$$\sum_{k=1}^{n-1}\sigma(S_{k-1},k)[W_k - W_{k-1}] \tag{9.35}$$

与被称为伊藤积分的随机变量

$$\int_0^T \sigma(S_u,u)\mathrm{d}W_u \tag{9.36}$$

差的方差在当 n 趋于无穷大时趋于0. 形式为

⊖ 右边的和式可以更详细地写成

$$\lim_{n\to\infty}\sum_{k=1}^{n-1}[a(S_{(k-1)h},kh)][(k)h-(k-1)h] \tag{9.30}$$

其中 $kh=t_k$.

$$\lim_{n\to\infty} E\Big[\sum_{k=1}^{n-1}\sigma(S_{k-1},k)[W_k - W_{k-1}] - \int_0^T \sigma(S_u,u)\mathrm{d}W_u\Big]^2 = 0 \tag{9.37}$$

9.2.3 定义：伊藤积分

现在我们可以结合随机微分方程来定义伊藤积分了.

定义 20 考虑随机微分方程在有限区间上的近似

$$S_k - S_{k-1} = a(S_{k-1},k)h + \sigma(S_{k-1},k)[W_k - W_{k-1}], \quad k = 1,2,\cdots,n \tag{9.38}$$

其中$[W_k - W_{k-1}]$是均值为 0 方差为 h 的标准维纳过程. 假设 $\sigma(S_t,t)$ 是非预期的，即它们与未来无关，且变量 $\sigma(S_t,t)$是非爆发性的，即：

$$E\Big[\int_0^T \sigma(S_t,t)^2 \mathrm{d}t\Big] < \infty \tag{9.39}$$

那么伊藤积分

$$\int_0^T \sigma(S_t,t)\mathrm{d}W_t \tag{9.40}$$

是当 $n\to\infty(h\to 0)$ 时，[㊀]

$$\sum_{k=1}^{n-1}\sigma(S_{k-1},k)[W_k - W_{k-1}] \to \int_0^T \sigma(S_u,u)\mathrm{d}W_u \tag{9.41}$$

的均方极限.

根据该定义，当区间数趋于无穷时每个区间的长度趋于 0，该有限和将趋于伊藤积分表示的随机变量. 显然该定义只有在极限随机变量存在时才有意义，可以证明 $\sigma(S_t,t)$非预期的假设是该极限存在的基本条件.[㊁]

简而言之，确定性积分与随机积分之间有 3 个主要差别：第一，随机积分中极限的定义不同；第二，伊藤积分只对非预期函数有定义；第三，标准微积分中使用函数的真实路径定义积分，而随机积分利用"随机等价"定义. 这些差别从本质上将随机积分同标准微积分区分开来.

下面的例子解释了在定义伊藤积分时为何要使用均方收敛，该例子还说明了伊藤积分不能按照路径定义的原因.

151

9.2.4 一个说明性的例子

伊藤积分是一个极限. 它是有限和的均方极限，因此，要使伊藤积分存在，求和式子必须收敛.

在适当的条件下，可以证明伊藤和收敛，且相应的伊藤积分存在. 然而通常难以确切地计算出均方极限的值，只有在某些特殊情况下例外. 本节我们考虑一个可以确切计算出

㊀ $[0,T]$被分割成了 n 个相等的区间，$T=nh$.

㊁ 一个关键点在于极限随机变量，即伊藤积分，是否依赖于$[0,T]$的分割方式，可以证明分割方式的选择不影响伊藤积分的值.

均方极限值的例子.[1]

假设我们要计算积分

$$\int_0^T x_t \mathrm{d}x_t \tag{9.42}$$

其中 $x_0=0$.

如果 x_t是确定性变量，我们可以用(9.24)中有限和的方法计算该积分. 首先将区间$[0,T]$分割成 n 个长度为 h 的子区间

$$t_0 = 0 < t_1 < \cdots < t_n \tag{9.43}$$

其中 $T=nh$，对任意 i，有 $t_{i+1}-t_i=h$.[2]其次，我们定义和

$$V_n = \sum_{i=0}^{n-1} x_{t_{i+1}} [x_{t_{i+1}} - x_{t_i}] \tag{9.44}$$

令 n 趋于无穷大. 其结果显而易见. 当 $x_0=0$ 时(9.42)的黎曼-斯蒂尔切斯积分为

$$\int_0^T x_t \mathrm{d}x_t = \frac{1}{2}x_T^2 \tag{9.45}$$

可以从图 9-1 中看出该积分的情况，其中我们考虑时间的任意函数 x_t，仅用一个矩形就可以得出曲线下方的面积.[3]

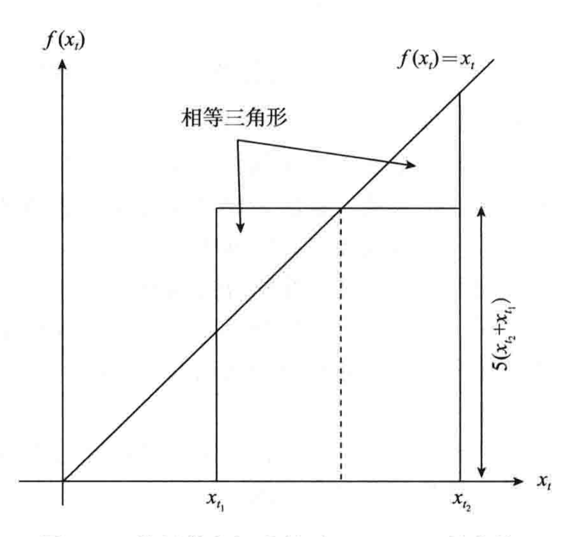


图 9-1 使用单个矩形得到 $f(x_t)=x_t$ 的曲线下方面积

如果 x_t 是维纳过程，就无法使用相同的方法. 首先要将 V_n调整为

$$V_n = \sum_{i=0}^{n-1} x_{t_i} [x_{t_{i+1}} - x_{t_i}] \tag{9.46}$$

也就是说，第一个 x_t 应该是时刻t_i 的值而不是时刻 t_{i+1}的值，否则这些项将无法保持不可预测性，因为在时刻 t_i 时 $x_{t_{i+1}}$ 是未知的，且与增量$[x_{t_{i+1}} - x_{t_i}]$ 相关. 在黎曼-斯蒂尔切斯积分中，使用两种类型的和最后得到的结果是相同的. 而在随机积分中，结果会因为取 $x_{t_{i+1}}$ 还是 x_{t_i}发生变化. 在后面我们将会看到，被积函数不可预测是伊藤积分存在的基本条件.

其次，V_n 现在是随机变量，没有简单的极限，可以通过概率方法得到 V_n 的极限，正如前面所言，伊藤积分使用的是均方极限.

因此我们要确定一个极限随机变量 V，使其满足

---

[1] 为与极限存在的证明进行对比.

[2] 等长子区间是为了方便，当 $t_{i+1}-t_i$ 不等时也可以得到相同的结果.

[3] 仅用一个矩形可行是因为被积函数 $f(x_t)$是一条倾斜角为 45 度的直线 $f(x_t)=x_t$.

$$\lim_{n\to\infty} E[V_n - V]^2 \to 0 \tag{9.47}$$

或等价地

$$\lim_{n\to\infty} E\Big[\sum_{i=0}^{n-1} x_{t_i}\Delta x_{t_i} - V\Big]^2 \to 0 \tag{9.48}$$

为了书写简洁，我们令

$$\Delta x_{t_{i+1}} = [x_{t_{i+1}} - x_{t_i}] \tag{9.49}$$

下面我们说明可以将该极限计算出来．　152

1. 均方极限的确切计算

在这一小节中，我们将通过一步步的计算得出极限随机变量 V，从而说明伊藤积分是随机变量和的均方极限的含义．第一步是确定 V_n 中的项．

对于任意 a，b，有

$$(a+b)^2 = a^2 + b^2 + 2ab \tag{9.50}$$

或

$$ab = \frac{1}{2}\Big[(a+b)^2 - a^2 - b^2\Big] \tag{9.51}$$

根据(9.46)，令 $a = x_{t_i}$，$b = \Delta x_{t_{i+1}}$，有

$$V_n = \frac{1}{2}\sum_{i=0}^{n-1}\Big[\Big(x_{t_i} + \Delta x_{t_{i+1}}\Big)^2 - x_{t_i}^2 - \Delta x_{t_{i+1}}^2\Big] \tag{9.52}$$

而

$$x_{t_i} + \Delta x_{t_{i+1}} = x_{t_{i+1}} \tag{9.53}$$

于是有

$$V_n = \frac{1}{2}\Big[\sum_{i=0}^{n-1} x_{t_{i+1}}^2 - \sum_{i=0}^{n-1} x_{t_i}^2 - \sum_{i=0}^{n-1}\Delta x_{t_{i+1}}^2\Big] \tag{9.54}$$

(9.54)中第一个和式与第二个和式除第一项和最后一项外相同，消掉这些项，且根据定义 $x_0 = 0$，⊖

$$V_n = \frac{1}{2}\Big[x_T^2 - \sum_{i=0}^{n-1}\Delta x_{t_{i+1}}^2\Big] \tag{9.55}$$

注意到 x_T 与 n 无关，所以 V_n 的均方极限等于 $\sum_{i=0}^{n-1}\Delta x_{t_{i+1}}^2$ 的均方极限．也就是说，我们现在需要找出一个 Z，使它满足下式

$$\lim_{n\to\infty} E\Big[\sum_{i=0}^{n-1}\Delta x_{t_{i+1}}^2 - Z\Big]^2 = 0 \tag{9.56}$$

153

该表达式左边有两个平方运算，一个是由于随机变量自身，另一个源于我们所使用的极限的类型．因此该极限中将含有 $\Delta x_{t_{i+1}}$ 的四阶项．

⊖ 因为令 $t_n = T$.

我们首先计算数学期望：

$$E\Big[\sum_{i=0}^{n-1}\Delta x_{t_{i+1}}^2\Big] \tag{9.57}$$

它看起来像是一个 Z 的备选值．直接求期望得

$$E\Big[\sum_{i=0}^{n-1}\Delta x_{t_{i+1}}^2\Big]=\sum_{i=0}^{n-1}E\Big[\Delta x_{t_{i+1}}^2\Big]=\sum_{i=0}^{n-1}\Big[t_{i+1}-t_i\Big] \tag{9.58}$$

可以简化为

$$\sum_{i=0}^{n-1}[t_{i+1}-t_i]=T \tag{9.59}$$

令它为 Z，我们可以计算期望

$$E\Big[\sum_{i=0}^{n-1}\Delta x_{t_{i+1}}^2-T\Big]^2=E\Big[\sum_{i=0}^{n-1}\Delta x_{t_{i+1}}^4+2\sum_{i=0}^{n-1}\sum_{j<i}^{n-1}\Big[\Delta x_{t_{i+1}}^2\Big]\times\Big[\Delta x_{t_{j+1}}^2\Big]+T^2-2T\sum_{i=0}^{n-1}\Big[\Delta x_{t_{i+1}}^2\Big]\Big] \tag{9.60}$$

现单独考虑(9.60)中右边各项，因为维纳过程的增量是相互独立的，所以

$$E\Big[\Delta x_{t_{i+1}}^2\Delta x_{t_{j+1}}^2\Big]=(t_{i+1}-t_i)(t_{j+1}-t_j) \tag{9.61}$$

且

$$E\Big[\Delta x_{t_{i+1}}^4\Big]=3(t_{i+1}-t_i)^2 \tag{9.62}$$

从而我们得到

$$E\Big[\sum_{i=0}^{n-1}\Delta x_{t_{i+1}}^2-T\Big]^2=\sum_{i=0}^{n-1}3(t_{i+1}-t_i)^2+2\sum_{i=0}^{n-1}\sum_{j<i}^{n-1}(t_{i+1}-t_i)(t_{j+1}-t_j)+T^2-2T\sum_{i=0}^{n-1}(t_{i+1}-t_i) \tag{9.63}$$

由于所有区间都被划分为相同的大小，对于任意 i 都有 $t_{i+1}-t_i=h$ ，因此我们有

$$\sum_{i=0}^{n-1}3(t_{i+1}-t_i)^2=3nh^2 \tag{9.64}$$

$$2\sum_{i=0}^{n-1}\sum_{j<i}^{n-1}(t_{i+1}-t_i)(t_{j+1}-t_j)=n(n-1)h^2 \tag{9.65}$$

且

$$T^2-2T\sum_{i=0}^{n-1}(t_{i+1}-t_i)=-T^2=-n^2h^2 \tag{9.66}$$

将它们结合到一起得到

$$E\Big[\sum_{i=0}^{n-1}\Delta x_{t_{i+1}}^2-T\Big]^2=3nh^2+n(n-1)h^2-n^2h^2 \tag{9.67}$$

即

$$E\left[\sum_{i=0}^{n-1}\Delta x_{t_{i+1}}^{2}-T\right]^{2}=2nh^{2}=2Th \tag{9.68}$$

因为当 $n\to\infty$ 时，区间的长度将趋于 0，所以

154

$$\lim_{h\to 0}E\left[\sum_{i=0}^{n-1}\Delta x_{t_{i+1}}^{2}-T\right]^{2}=\lim_{h\to 0}2hT=0 \tag{9.69}$$

因此，$\sum_{i=0}^{n-1}\Delta x_{t_{i+1}}^{2}$ 的均方极限是 T.

现回到 V_n，

$$V_n=\frac{1}{2}\left[x_T^2-\sum_{i=0}^{n-1}\Delta x_{t_{i+1}}^{2}\right] \tag{9.70}$$

我们发现 V_n的均方极限为

$$\lim_{n\to\infty}E[V_n]^2=\frac{1}{2}\left[x_T^2-T\right] \tag{9.71}$$

右边项是伊藤积分

$$\int_0^T x_t\,\mathrm{d}x_t \tag{9.72}$$

可以看得伊藤积分的表达式不同于标准微积分，伊藤积分值为

$$\int_0^T x_t\,\mathrm{d}x_t=\frac{1}{2}\left[x_T^2-T\right] \tag{9.73}$$

而黎曼积分中是没有 T 的那一项的.

这是利用均方极限确切地计算出伊藤积分的一个例子，其中伊藤积分为极限随机变量 $\frac{1}{2}\left[x_T^2-T\right]$.

2. 一个重要的评注

在前一节中

$$\lim_{h\to 0}E\left[\sum_{i=0}^{n-1}\Delta x_{t_i+1}^{2}-T\right]^{2}=0 \tag{9.74}$$

我们感兴趣的是将它写成积分形式.

假设 x_t 是维纳过程，考虑积分

$$\int_0^T(\mathrm{d}x_t)^2 \tag{9.75}$$

可以把它理解为是 x_t 增量的平方和.

如果该积分在伊藤意义下存在，那么根据定义

$$\lim_{h\to 0}E\left[\sum_{i=0}^{n-1}\Delta x_{t_i+1}^{2}-\int_0^T(\mathrm{d}x_t)^2\right]^{2}=0 \tag{9.76}$$

但是我们知道

$$\int_0^T\mathrm{d}t=T \tag{9.77}$$

将等式(9.74)(9.76)和(9.77)结合起来，会得到一个与标准微积分相比看起来有些不同寻常的结果：

$$\int_0^T (\mathrm{d}x_t)^2 = \int_0^T \mathrm{d}t \tag{9.78}$$

该等式是在均方极限的意义下成立的．因此，如果 W_t 表示维纳过程，对于任意无穷小的 $\mathrm{d}t$，都有

$$\mathrm{d}W_t^2 = \mathrm{d}t \tag{9.79}$$

事实上，在所有关于随机微积分的实际运算中，将含有 $\mathrm{d}W_t^2$ 的项用 $\mathrm{d}t$ 替换是非常普遍的，前面的讨论给出了这一替换的背后原因，该等式仅在均方收敛意义下才成立．

9.3 伊藤积分的性质

考虑随机微分方程

155

$$\mathrm{d}S_t = a(S_t, t)\mathrm{d}t + \sigma(S_t, t)\mathrm{d}W_t \tag{9.80}$$

在区间[0, T]上对该等式积分

$$\int_0^T \mathrm{d}S_t = \int_0^T a(S_t, t)\mathrm{d}t + \int_0^T \sigma(S_t, t)\mathrm{d}W_t \tag{9.81}$$

右边第二项积分是伊藤意义下的积分，那么该积分有什么性质呢？

9.3.1 伊藤积分是鞅

可以证明伊藤积分是鞅，这一性质在金融理论和实际资产价格计算中对于资产价格干扰项的建模非常有用．

描述资产价格动态变化的模型含有代表不可预测信息的干扰项．下列形式的积分⊖

$$\int_t^{t+\Delta} \sigma_u \mathrm{d}W_u \tag{9.82}$$

是在长度为 Δ 的区间内资产价格不可预测的扰动和．如果已知 t 时刻信息时任意一个增量都是不可预测的，那么这些增量的和也是不可预测的，这使得(9.82)中的积分为*鞅差分*：

$$E\left[\int_t^{t+\Delta} \sigma_u \mathrm{d}W_u\right] = 0 \tag{9.83}$$

那么积分

$$\int_0^t \sigma_u \mathrm{d}W_u \tag{9.84}$$

就是鞅：

$$E_s\left[\int_t^{t+\Delta} \sigma_u \mathrm{d}W_u\right] = \int_0^t \sigma_u \mathrm{d}W_u \tag{9.85}$$

因此，描述资产价格动态变化的等式中不可预测干扰项的存在刚好符合了伊藤积分的鞅性

⊖ 为简化书写，令 $\sigma(S_u, u) = \sigma_u$.

质，要求 σ_t 在已知信息集 I_t 时是不可预测的这一条件保证了伊藤积分的鞅性质．

1. **情况 1**

假设波动率参数 $\sigma(S_t,t)$ 是独立于资产价格 S_t 和时间 t 的常数：

$$\sigma(S_t,t)=\sigma \tag{9.86}$$

那么伊藤积分将等于黎曼积分，为

$$\int_t^{t+\Delta}\sigma \mathrm{d}W_t=\sigma[W_{t+\Delta}-W_t] \tag{9.87}$$

考虑该积分的预测

$$E_s\left[\int_t^{t+\Delta}\sigma \mathrm{d}W_u \,\middle|\, \int_0^t \sigma \mathrm{d}W_u\right]=\int_0^t \sigma \mathrm{d}W_u=\sigma(W_t-W_0) \tag{9.88}$$

其中 $\Delta>0$．这是一种特殊情况，因为维纳过程的增量均值为 0 且互不相关：

$$E_s[\sigma(W_{t+\Delta}-W_0)\,|\,(W_{t+\Delta}-W_0)] \tag{9.89}$$

$$=E_s[\sigma(W_{t+\Delta}-W_t)+\sigma(W_t-W_0)\,|\,(W_{t+\Delta}-W_0)] \tag{9.90}$$

$$=\sigma(W_t-W_0) \tag{9.91}$$

我们可以再次看出伊藤积分具有鞅性质．[㊀]

因此当 σ 是常数时，黎曼积分和伊藤积分刚好相等，且都是鞅．

2. **情况 2**

另一方面，如果 σ 的取值依赖于 S_t，而 S_t 依赖于 W_t，那么伊藤积分将不同于黎曼积分，但伊藤积分仍是鞅，而黎曼积分不再是鞅． 156

例如，如果标的资产价格服从几何分布，且扩散项中

$$\sigma(S_t,t)=\sigma S_t \tag{9.92}$$

那么伊藤积分将不同于黎曼积分，用黎曼和来对伊藤积分进行近似将导致自相矛盾.

我们用下面的例子说明这一点．

3. **一个例子**

假设资产价格服从随机微分方程

$$\mathrm{d}S_t=a(S_t,t)\mathrm{d}t+\sigma(S_t,t)\mathrm{d}W_t,\quad 0\leqslant t \tag{9.93}$$

其中漂移系数和扩散系数为

$$a(S_t,t)=\mu S_t \tag{9.94}$$

$$\sigma(S_t,t)=\sigma S_t \tag{9.95}$$

两个系数都与资产价格 S_t 成正比．

再次考虑长度为 Δ 的小区间，并在其上积分：

$$\int_t^{t+\Delta}\mathrm{d}S_u=\int_t^{t+\Delta}\mu S_u \mathrm{d}u+\int_t^{t+\Delta}\sigma S_u \mathrm{d}W_u \tag{9.96}$$

项 $\sigma S(W_t)$ 通过 S_t 与 W_t 间接相关．[㊁]

㊀ 记住 $W_0=0$．

㊁ 这里我们使用 $S(W_t)=S_t$ 来简化解释．

现在考虑，如果我们用黎曼和对右边第二个积分项进行近似会有什么结果．

用黎曼和近似需要用到子区间中点处观察到的维纳过程的值．这相当于先计算

$$\sigma S\left(\frac{W_{t+\Delta}+W_t}{2}\right) \tag{9.97}$$

再用其乘以矩形的底 $W_{t+\Delta}-W_t$．

黎曼和将含有如下项

$$\sigma S\left(\frac{W_{t+\Delta}+W_t}{2}\right)\left[W_{t+\Delta}-W_t\right] \tag{9.98}$$

显然，由于 $S(\cdot)$及矩形底中含有相互关联的项，这些项的数学期望不是 0.

考虑随机微分方程如下的简单情况

$$\mathrm{d}S_t=\sigma W_t\mathrm{d}W_t \tag{9.99}$$

方程中干扰项为

$$\int_t^{t+\Delta}\sigma W_u\mathrm{d}W_u \tag{9.100}$$

为了使用黎曼和近似该积分项，我们使用底为 $W_{t+\Delta}-W_t$、高为 $\sigma\left(\frac{W_{t+\Delta}+W_t}{2}\right)$ 的矩形：[⊖]

$$\int_t^{t+\Delta}\sigma W_u\mathrm{d}W_u\approx\left[\sigma\frac{W_{t+\Delta}+W_t}{2}\right](W_{t+\Delta}-W_t) \tag{9.101}$$

然而，对右边项进行条件期望运算 $E[\cdot]$，

$$E\left[\left(\frac{W_{t+\Delta}+W_t}{2}\right)(W_{t+\Delta}-W_t)\,\middle|\,W_t\right]=E\left[\frac{1}{2}\left(W_{t+\Delta}^2+W_t^2\right)\,\middle|\,W_t\right] \tag{9.102}$$

$$=\frac{1}{2}\Delta \tag{9.103}$$

且 $\Delta\neq0$，也就是说近似和的数学期望不为 0，它是预测的，显然这会与左边是干扰项的积分矛盾.

157

如果这些相关系数不为 0，利用黎曼和来近似伊藤积分会使干扰项的数学期望非零：

$$E\left[\int_t^{t+\Delta}\sigma_s\mathrm{d}W_s\right]\neq0,\quad 0<\Delta \tag{9.104}$$

为了保留 $\sigma(S_t,t)$非预报的性质，伊藤积分近似必须用如下矩形

$$\sigma(S_t,t)(W_{t+\Delta}-W_t) \tag{9.105}$$

其中根据定义 $\sigma(S_t,t)$与增量 ΔW_t 不相关．

上述讨论表明黎曼积分与资产定价模型中的假设不一致，除了如下特殊情况

$$\sigma(S_t,t)=\sigma(t) \tag{9.106}$$

此外，如果被积函数不是不可预测的，将无法保证伊藤积分中的部分和均方收敛于有意义的随机变量．除此之外，相较于失去鞅性质，还存在一个更基本的问题：积分可能不

⊖ 为了简洁我们只使用了一个矩形，事实上是可以对区间$[t,t+\Delta]$进行更精细的分割的．

存在．

下一节将对这一点进行简要讨论．

9.3.2 路径积分

在随机微积分中经常会出现随机积分无法点态定义的情况．这意味着什么呢？

考虑二点过程 $S_{t_{i+1}}-S_{t_i}$（$i=1, 2, \cdots, n$），它在$[0,T]$上长度为 Δ 的离散区间上取值：

$$S_{t_{i+1}}-S_{t_i}=\begin{cases}\sqrt{\Delta} & \text{概率为 } p\\ -\sqrt{\Delta} & \text{概率为 } 1-p\end{cases}\tag{9.107}$$

与前面一样，$T=n\Delta$.

该过程的典型路径是一系列$\sqrt{\Delta}$和$-\sqrt{\Delta}$交互出现．例如，一个典型路径可能如下所示：

$$\{\sqrt{\Delta},\sqrt{\Delta},-\sqrt{\Delta},\sqrt{\Delta},\cdots\}\tag{9.108}$$

假设一名金融分析师需要对

$$\int_0^T f(S_t)\mathrm{d}S_t$$

进行近似，会用到如下的有限和：

$$V_n=\sum_{i=0}^{n-1}f(S_{t_{i+1}})\Big[S_{t_{i+1}}-S_{t_i}\Big]\tag{9.109}$$

假设利用 S_t 的特定路径计算 V_n. 例如，考虑正负交替出现的路径：

$$\{\sqrt{\Delta},-\sqrt{\Delta},\sqrt{\Delta},-\sqrt{\Delta},\cdots,\sqrt{\Delta}\}\tag{9.110}$$

用这些观测值替代 V_n 中的 $S_{t_{i+1}}-S_{t_i}$ 可以得到

$$V_n=\Big[f(\sqrt{\Delta})(\sqrt{\Delta})+f(0)(-\sqrt{\Delta})+f(\sqrt{\Delta})(\sqrt{\Delta})+\cdots\Big]\tag{9.111}$$

V_n 的值依赖于 S_t 的特定轨迹．如果 V_n 收敛，那么我们称之为点态积分．

可以证明在随机环境下这种点态积分不一定收敛．考虑下面一个简单的例子．

令 V_n 中函数 $f(\cdot)$为

$$f(S_{t_{i+1}})=\operatorname{sign}(S_{t_{i+1}}-S_{t_i})\tag{9.112}$$

也就是说我们根据 $S_{t_{i+1}}-S_{t_i}$ 的符号确定 $f(\cdot)$的值为$+1$ 还是-1.

这意味着 V_n 中所有项都大于零，因此

$$V_n=\sum_{i=0}^{n-1}\sqrt{\Delta}=n\sqrt{\Delta}\tag{9.113}$$

158

利用 $T=n\Delta$，

$$V_n=\frac{T}{\sqrt{\Delta}}\tag{9.114}$$

显然当 $\Delta\to 0$ 时，V_n 趋于无穷大．

如果这类路径出现的概率大于零，那么路径和 V_n 在任何概率意义下都不收敛．

出于两个原因，本例非常重要：

首先，我们可以从这个例子中看出路径依赖积分的含义．在计算点态积分时，我们没有用到与 $\Delta S_{t_{i+1}}$ 相关的概率，该积分是利用随机过程的真实概率进行计算的，而伊藤积分是利用均方收敛进行计算的，且积分是在随机等价环境中确定的．

其次，我们可以看出利用非预报函数 $f(\cdot)$的重要性．事实上，如果可以预测未来的 $f(\cdot)$，那么就可以预测 $S_{t_{i+1}}-S_{t_i}$ 的符号，这使和式中所有项都大于零，并会导致 V_n 在 n 变大时急剧增加．

9.3.3 伊藤等距

任意平方可积的 $X_t(\omega)$都有

$$E\left(\left(\int_0^t X_u(\omega)\mathrm{d}W_u(\omega)\right)^2\right)=E\left(\int_0^t X_u(\omega)^2\mathrm{d}u\right) \tag{9.115}$$

这一性质称为伊藤等距．

伊藤等距的一个例子

称下面的随机微分方程为 Ornstein- Uhlenbeck 方程

$$\mathrm{d}X_t = \mu X_t\mathrm{d}t+\sigma\mathrm{d}B_t \tag{9.116}$$

其中 μ 和 σ 是实常数．我们可以用伊藤引理来求解，为此我们需要使用积分系数 $\mathrm{e}^{-\mu t}$，然后对 $\mathrm{e}^{-\mu t}X_t$ 使用伊藤引理，计算 $\mathrm{d}(\mathrm{e}^{-\mu t}X_t)$

$$\mathrm{d}(\mathrm{e}^{-\mu t}X_t)=(-\mu\mathrm{e}^{-\mu t}X_t+\mathrm{e}^{-\mu t}X_t)\mathrm{d}t+\mathrm{e}^{-\mu t}\sigma\mathrm{d}B_t \tag{9.117}$$

$$=\mathrm{e}^{-\mu t}\sigma\mathrm{d}B_t \tag{9.118}$$

然后对它求积分，得到

$$\mathrm{e}^{-\mu t}X_t = X_0+\int_0^t \mathrm{e}^{-\mu s}\sigma\mathrm{d}B_s \tag{9.119}$$

或者等价地

$$X_t = \mathrm{e}^{\mu t}X_0+\int_0^t \mathrm{e}^{\mu(t-s)}\sigma\mathrm{d}B_s \tag{9.120}$$

对 X_t 进行处理后，我们现在可以得出其均值 $E(X_t)$和方差 $\mathrm{Var}(X_t)$.

由于 $\int_0^t \mathrm{e}^{\mu(t-s)}\mathrm{d}B_s$ 是随机积分，那么 $\int_0^t \mathrm{e}^{\mu(t-s)}\mathrm{d}B_s$ 也是鞅，因此

$$E(X_t) = \mathrm{e}^{\mu t}X_0+0 = \mathrm{e}^{\mu t}X_0 \tag{9.121}$$

并且

$$E(X_t^2)=E\left(\mathrm{e}^{2\mu t}X_0^2+2\mathrm{e}^{\mu t}X_0\int_0^t \mathrm{e}^{\mu(t-s)}\sigma\mathrm{d}B_s+\left(\int_0^t \mathrm{e}^{\mu(t-s)}\sigma\mathrm{d}B_s\right)^2\right) \tag{9.122}$$

$$=\mathrm{e}^{2\mu t}X_0^2+2\mathrm{e}^{\mu t}X_0E\left(\int_0^t \mathrm{e}^{\mu(t-s)}\sigma\mathrm{d}B_s\right)+E\left(\left(\int_0^t \mathrm{e}^{\mu(t-s)}\sigma\mathrm{d}B_s\right)^2\right) \tag{9.123}$$

该表达式中第二项是随机积分，因此它是鞅．对最后一项应用伊藤等距，得到

$$E(X_t^2)=\mathrm{e}^{2\mu t}X_0^2+0+E\left(\int_0^t \mathrm{e}^{2\mu(t-s)}\sigma^2\mathrm{d}t\right) \tag{9.124}$$

$$=e^{2\mu t}X_0^2+\int_0^t e^{2\mu(t-s)}\sigma^2\,dt \tag{9.125}$$

$$=e^{2\mu t}X_0^2+\frac{\sigma^2}{2\mu}(e^{2\mu t}-1) \tag{9.126}$$

159

因此

$$\mathrm{Var}(X_t)=E(X_t^2)-(E(X_t))^2 \tag{9.127}$$

$$=\frac{\sigma^2}{2\mu}(e^{2\mu t}-1) \tag{9.128}$$

9.4　伊藤积分的其他性质

伊藤积分有一些其他性质.

9.4.1　存在性

我们可能会提出这样一个问题：对于(9.6)给出的$\{S_t\}$，一般随机函数 $f(S_t, t)$的伊藤积分什么时候存在？

$$\int_0^t f(S_u,u)\,dS_u \tag{9.129}$$

可以证明：如果函数 $f(\cdot)$连续且是非预报的，该积分存在．也就是说有限和

$$\sum_{i=0}^{n-1} f(S_{t_i},t_i)[S_{t_{i+1}}-S_{t_i}] \tag{9.130}$$

均方收敛于某个被称为伊藤积分的随机变量.⊖

9.4.2　相关性

因为伊藤积分是一个随机变量(更准确地说是一个随机过程)，因此它包含了很多不同时刻的值.

鞅性质给出了不可预测函数 $f(\cdot)$关于维纳过程积分的一阶矩的值

$$E\left[\int_0^T f(W_t,t)\,dW_t\right]=0 \tag{9.131}$$

其中 W_t 是维纳过程．二阶矩的方差和协方差分别为

$$E\left[\int_0^T f(W_t,t)\,dW_t\int_0^T g(W_t,t)\,dW_t\right]=\int_0^T E[f(W_t,t)g(W_t,t)]\,dt \tag{9.132}$$

与

$$E\left[\int_0^T f(W_t,t)\,dW_t\right]^2=E\left[\int_0^T f(W_t,t)^2\,dt\right] \tag{9.133}$$

推导中用到了等价式 $dW_t^2=dt$.

⊖ 尽管它可能存在，但不能直接确定这样一个极限.

9.4.3 可加性

伊藤积分也具有一些类似于黎曼-斯蒂尔切斯积分的性质.

特别地，(9.6)中两个关于 S_t 的(随机)函数和的积分等于它们积分的和：

$$\int_0^T [f(S_t,t)+g(S_t,t)]\mathrm{d}S_t = \int_0^T f(S_t,t)\mathrm{d}S_t + \int_0^T g(S_t,t)\mathrm{d}S_t \tag{9.134}$$

9.5 关于带跳过程的积分

极端不规律的连续时间鞅和维纳过程使随机积分的定义变得更加复杂，这使得点态积分变为不可能.

那么对于跳跃过程进行随机积分也会出现相同的问题吗？我们可以在处理泊松过程时160使用黎曼-斯蒂尔切斯积分吗？

令人惊讶的是，这个问题的答案在某些条件下是肯定的.

假设 M_t 是鞅，存在有限个跳跃且不含维纳成分，M_t 的轨迹除个别跳跃外非常光滑. 我们可以定义一个点态的 V_n

$$V_n = \sum_{i=0}^{n-1} f(M_{t_i})[M_{t_{i+1}} - M_{t_i}] \tag{9.135}$$

该 V_n 是收敛的，且随机过程 M_t 变化量有限的概率为 1，在这种情况下，我们可以认为 V_n 是点态收敛的.

9.6 小结

本章对伊藤积分进行了定义.

对于从业者来说要记住，有两点非常重要：第一，随机微分方程中的误差项是定义在伊藤积分意义下的，数值计算时必须遵循该定义的条件；第二，资产定价中用到的随机微分方程也是定义在伊藤积分意义下的.

最重要的是，伊藤积分是一些随机和的均方极限，我们将这些随机和小心地组合在一起，以确保积分结果是鞅.

我们也对一些例子进行了讨论，并说明了与确定性环境相比随机环境下的积分法则非常不同，这是使用均方极限的结果.

幸运的是，在求伊藤积分时很少会用到直接求均方极限的方法，而是采用更直接的伊藤引理. 我们将在下一章中讨论伊藤引理，并对伊藤积分进行更加深入的讨论.

9.7 参考阅读

在伊藤积分的导数方面有很多非常优秀的书籍，比如，我们已经提到过的 Karatzas 和 Shreve(1991)，Revuz 和 Yor(1994)，读者可能会发现 Oksendal(2000)和 Protter(1990)读起来更加容易. 对于量化方向的从业者和刚刚开始学习的研究生来说，Oksendal(2000)可

以作为一本很好的参考手册，它写得非常好并且易于理解，术语很少.

9.8 习题

1. W_t 是定义在$[0,T]$上的维纳过程，考虑积分

$$\int_0^t W_t^2 \mathrm{d}W_s$$

在接下来的问题中使用$[0,t]$的子区间：

$$t_0, t_1, \cdots, t_{n-1}, t_n$$

(a) 将上面的积分写成 3 个不同黎曼和的近似式.

(b) 用伊藤和写出该积分的离散时间形式.

(c) 计算 3 个黎曼和的数学期望.

(d) 计算伊藤和的数学期望.

161

2. 证明：已知 t_0，t_1，…，t_{n-1}，t_n 和 W_{t_0}，W_{t_1}，…，$W_{t_{n-1}}$，W_{t_n}，总有：

$$\sum_{j=1}^{n}\left[t_j W_{t_j} - t_{j-1} W_{t_{j-1}}\right] = \sum_{j=1}^{n}\left[t_j\left(W_{t_j} - W_{t_{j-1}}\right)\right] + \sum_{j=1}^{n}\left[\left(t_j - t_{j-1}\right)W_{t_j}\right] \tag{9.136}$$

它与普通函数乘积的微分法则

$$\mathrm{d}(uv) = (\mathrm{d}u)v + u(\mathrm{d}v)$$

有什么不同?

3. 使用上题结论证明：

$$\int_0^t s\,\mathrm{d}W_s = tW_t - \int W_s \mathrm{d}s$$

4. 上式中有两个积分项，其中哪一项仅在伊藤积分的定义下有意义?

5. 我们可以认为这是变量代换吗?

6. 我们可以认为这是分部积分吗?

7. Y 为随机变量，且定义为

$$Y(\omega) = \int_0^1 W_s(\omega)\mathrm{d}s \tag{9.137}$$

其中 W 是标准维纳过程，计算 $E(Y)$和 $E(Y^2)$.

8. 编写一个可以通过拟合来对如下随机积分进行数值计算的程序：

$$X_t = \int_0^t W_t^2 \mathrm{d}W_s \tag{9.138}$$

9. 考虑随机过程 X_t

$$X_t = \int_0^t \gamma_s \mathrm{d}B_s \tag{9.139}$$

其中 γ_s 是随机的，B_s 是标准布朗运动. 写出当 $s<t$ 时 $\mathrm{Cov}(X_s, X_t)$的表达式(应用伊藤等距).

10. 证明

$$\int_0^t B_s \mathrm{d}B_s = \frac{1}{2}B_t^2 - \frac{1}{2}t$$

观察项$-\frac{1}{2}t$，证明随机积分的结果与确定性微积分的结果形式不同，而且含义不同.

162

第 10 章 伊藤引理

10.1 引言

前面已经讨论过，在随机环境下按常规方式定义导数可能不存在．假设资产价格的波动不可预测，在连续时间环境下它们会变得过于不稳定，得到的资产价格可能是连续的，但不一定光滑，因此需要用随机微分来替代普通导数．

伊藤法则为简化随机微分计算并求出具体的值提供了一种工具，这是本章的主要内容．

在本章的开始部分，我们将介绍几种不同的导数．

10.2 导数的类型

假设 $F(S_t, t)$是关于变量 S_t、t 的函数，其中 S_t 本身随时间 t 变化．进一步假设 S_t 是随机过程．

在标准微积分中所有变量都是确定性的，存在三种类型的导数．

第一种是 $F(S_t, t)$的偏导数，记为

$$F_s = \frac{\partial F(S_t, t)}{\partial S_t}, \quad F_t = \frac{\partial F(S_t, t)}{\partial t} \tag{10.1}$$

第二种是微分形式的全导数：

$$\mathrm{d}F_t = F_s \mathrm{d}S_t + F_t \mathrm{d}t \tag{10.2}$$

在(10.2)中，$\mathrm{d}F_t$ 是 $\mathrm{d}F(S_t, t)$的简写，不要将它与 $F(\cdot)$关于 t 的偏导数 F_t 相混淆．

第三种是链式法则：

$$\frac{\mathrm{d}F(S_t, t)}{\mathrm{d}t} = F_s \frac{\mathrm{d}S_t}{\mathrm{d}t} + F_t \tag{10.3}$$

金融市场参与者可能出于各种原因对这三种导数感兴趣．

偏导数在真实世界中除了可以在风险方面给出资产价格对观察变化的反应系数外，没有直接的应用．例如，F_s 仅度量了 $F(S_t, t)$对于 S_t 的微小变化做出的反应，就其本身而言，由于连续随机变量 S_t 只有在时间发生变化时才能改变，所以 F_s 只是一个假想的概念，因为实际上 t 也要变．偏导数从这种问题中抽象而来，由于它们都是简单系数，因此随机环境和确定性环境中定义偏导数的方式没有差别．

一个使用偏导数的经典例子就是 delta 对冲．假设市场参与者知道 $F(S_t, t)$的函数形式．为了得到偏导数 F_s，该数学公式只对 S_t 求导．这里 F_s 度量的是衍生资产价格在 S_t 每变化一单位时所发生的变化．在该意义下，不会出现定义维纳过程的时间导数时遇到的问题．我们检验的不是 $F(S_t, t)$如何随时间变化，而是 $F(\cdot)$在假设时间固定的条件下如何对 S_t 发生的微小变化进行反应．

全导数是一个更具有现实意义的概念．它假设时间 t 和标的证券价格 S_t 都发生变化，计算 $F(S_t,t)$的整体反应，这就需要(随机)微分 dF_t. 对于市场参与者来说，这显然是一个非常实用的量．它代表了在区间 dt 内观察到的衍生资产价格变化．

链式法则与全导数非常相像．在经典微积分中，链式法则表示某个自变量的微小变化对整个函数的链式影响．我们知道在随机微积分中，连续时间的平方可积鞅或者布朗运动是无法定义类似于 dF_t/dt、dS_t/dt 这样的运算的，但可以得到用 dF_t、dS_t、dt 的绝对变化量给出的链式法则，而这就需要用伊藤积分给出．因此将“链式法则”运用于随机微积分中，也就是随机形式的全微分．

例子

在介绍伊藤公式前，我们先来讨论一个简单的例子，这个例子会帮助我们区分取不同导数的技巧．$F(r_t,t)$为 T 时刻到期的短期国债价格，r_t 为固定的无风险连续复利率．于是有

$$F(r_t,t) = e^{-r_t(T-t)}100 \tag{10.4}$$

计算偏导数 F_r、F_t：

$$F_r = \frac{\partial F}{\partial r_t} = -(T-t)[e^{-r_t(T-t)}100] \tag{10.5}$$

164

$$F_t = \frac{\partial F}{\partial t} = r_t[e^{-r_t(T-t)}100] \tag{10.6}$$

注意无论 r_t 是确定的还是随机的，这些偏导数都是如此．通过计算偏导数，我们得出了 $F(\cdot)$在 r_t 或 t 发生假想的微小变化时的变化率．

另一方面，全导数对应的是真实发生的随机事件．在标准微积分中，r_t 非随机，该特殊的 $F(\cdot)$的全导数为

$$dF(r_t,t) = -(T-t)[e^{-r_t(T-t)}100]dr_t + r_t[e^{-r_t(T-t)}100]dt \tag{10.7}$$

通过这个例子我们可以设想，当 r_t 是随机变量时，我们也许能够利用伊藤积分定义随机函数的全导数，因为伊藤积分给出了 dr_t 等随机微分的意义，这种直觉是正确的，得到的结果就是伊藤公式．但是当 r_t 是随机变量时，不仅 dr_t 的含义发生了变化，[㊀] 全导数的公式也会变得不同.

10.3 伊藤引理

伊藤引理是随机形式的链式法则．设 S_t 是一个依赖于维纳过程 W_t 的连续时间的随机过程，且 $F(S_t,t)$是一个包含 S_t 的已知函数．我们要计算的是当时间有一个增量变化 dt 时，$F(\cdot)$发生的变化．显然，时间推移会在两方面影响 $F(S_t,t)$. 首先会有通过 $F(S_t,t)$ 中 t 变量带来的直接影响；其次，随着时间的变化，我们可以得到关于 W_t 的新信息，观察到

㊀ 这种量是定义在均方收敛和随机等价的意义下的．

新的增量 dS_t，这也会使得 $F(\cdot)$发生变化．这两种影响的和我们用链式法则的随机微分形式 $dF(S_t,t)$给出．

在连续时间下观察随机过程 S_t，我们再次将时间区间$[0,T]$划分为 n 个长度均为 h 的等长小区间，利用有限差分近似表示．等式

$$\Delta S_k = a_k h + \sigma_k \Delta W_k, \quad k = 1,2,\cdots,n \tag{10.8}$$

用到了当 $h\to 0$ 时左右两边均方等价．这个记法将贯穿本章始终，其中，a_k、σ_k 分别是 $a(S_{k-1},k)$、$\sigma(S_{k-1},k)$的简写．

在这样的背景下，我们利用泰勒级数来推导伊藤公式．回顾光滑（即无穷次可积）函数 $f(x)$在任意点 x_0 附近的泰勒级数展开：

$$f(x) = f(x_0) + f_x(x_0)(x-x_0) + \frac{1}{2}f_{xx}(x_0)(x-x_0)^2 + R \tag{10.9}$$

其中 R 表示余项．

对 $F(S_t,t)$应用该公式．(10.9)式要求 $F(\cdot)$是 S_t 的光滑函数．[⊖]但现在有两点使问题变复杂了．首先，(10.9)中的泰勒级数公式适用于单变量函数 $f(x)$，而 $F(S_t,t)$依赖于 S_t 和 t 两个变量．其次，(10.9)对于确定性变量有效，而 S_t 是随机过程．在应用泰勒级数前要先处理这些复杂问题．

将单变量泰勒级数推广到双变量情况非常简单．添加关于第二个变量的偏导数即可．对于双变量的情形，交叉偏导数也必须添加进去．

165

泰勒级数在随机环境中的应用还有一些更深层次的问题．首先，要记得泰勒级数中的一些项是偏导数，关于这一点，在随机环境中进行微分没有难度．其次，会出现 dS_t 这种微分．对这种项，我们需要对等式的意义进行一些调整，泰勒级数展开本身不需要变动，泰勒级数的公式也保持不变，但等式的意义要发生改变，这种等式是建立在均方收敛的意义下的．

对 $F(S_k,k)(k=1, 2, \cdots)$进行泰勒展开，并假设其中的 S_k 服从

$$\Delta S_k = a_k h + \sigma_k \Delta W_k \tag{10.10}$$

首先固定 k，给定信息集 I_{k-1}后，S_{k-1}就是一个已知值．然后用泰勒公式在 S_{k-1}和 $k-1$ 附近展开 $F(S_k,k)$：

$$F(S_k,k) = F(S_{k-1},k-1) + F_s[S_k - S_{k-1}] + F_t[h] + \frac{1}{2}F_{ss}[S_k - S_{k-1}]^2$$

$$+ \frac{1}{2}F_{tt}[h] + F_{st}[h(S_k - S_{k-1})] + R \tag{10.11}$$

为了方便，我们使用了 F_t，F_{st}，F_{tt} 这些记号，虽然它们是关于 k 的偏导数，但公式中这些的偏导数 F_s，F_{ss}，F_t，F_{tt}，F_{st}都是在 S_{k-1}，$k-1$ 处取值，R 表示泰勒级数的余项．

将(10.11)中 $F(S_{k-1},k-1)$移到等式左边并重新标注，有：

$$F(S_k,k) - F(S_{k-1},k-1) = \Delta F(k) \tag{10.12}$$

⊖ 有些读者可能想知道这种光滑性是否与 S_t 的极端不规则性相冲突．$F(\cdot)$可以同时为 S_t 的光滑函数和非常不规则的随机过程．这里的不规则性指的是 $F(\cdot)$如何随时间变化，而不是说明 S_t 如何与 $F(\cdot)$相关．

$$S_k - S_{k-1} = \Delta S_k \tag{10.13}$$

注意，等式(10.11)对于时间变量使用了增量：

$$kh - (k-1)h = h \tag{10.14}$$

现在将它们带入(10.11)：

$$\Delta F(k) = F_s \Delta S_k + F_t[h] + \frac{1}{2}F_{ss}[\Delta S_k]^2 + \frac{1}{2}F_{tt}[h]^2 + F_{st}[h\Delta S_k] + R \tag{10.15}$$

因为 S_t 的变化由(10.10)给出，即

$$\Delta S_k = a_k h + \sigma_k \Delta W_k \tag{10.16}$$

我们可以用该等式右边的项替代(10.11)泰勒级数展开中的 ΔS_k，则有

$$\begin{aligned}\Delta F(k) =& F_s[a_k h + \sigma_k \Delta W_k] + F_t[h] + \frac{1}{2}F_{ss}[a_k h + \sigma_k \Delta W_k]^2 \\ &+ \frac{1}{2}F_{tt}[h]^2 + F_{st}[h(a_k h + \sigma_k \Delta W_k)] + R\end{aligned} \tag{10.17}$$

这个等式是什么意思呢？等式左边 $\Delta F(k)$表示 $F(S_k, k)$由 k 和 S_k 变动产生的总变化量．因此，如果 $F(S_k, k)$表示衍生证券价格，左边得到的就是在小区间内衍生资产价格的变化量．该变化量由右边项解释．

一阶项 $F_t[h]$表示时间变化带来的影响，$F_s[a_k h + \sigma_k \Delta W_k]$表示标的证券价格变化带来的影响，在后一项中，我们再次看到证券价格的变化同时包含可预测部分和不可预测部分．二阶项为平方项和交叉乘积项的变化量．高阶项全都包含在余项 R 中．

为了得到随机环境下的链式法则，可以将右边项分为可忽略项和不可忽略项．接下来将证明在微小时间区间内可忽略项可以从右边舍去，得到链式法则．此外，当 $h \to 0$ 时，可以进行极限的讨论，并得到一个均方意义下精确的公式．该公式就是伊藤引理．

166

这种方法的第一步是将右边项进行分离，这需要一个确切的标准来决定哪些项可被忽略．之后分别考虑(10.11)右边项的大小，决定哪些项可以舍掉．

10.3.1 随机微积分中“大小”的概念

本节介绍用于确定随机微积分中哪些项可忽略的常用方法．

在标准微积分中，函数 $f(S)$在 S_0 附近的泰勒级数展开为

$$f(S) - f(S_0) = \Delta f = f_s(S_0)\Delta S + \frac{1}{2}f_{ss}(S_0)\Delta S^2 + \frac{1}{3}f_{sss}(S_0)\Delta S^3 \tag{10.18}$$

其中 R 是余项．而全导数的公式为

$$\mathrm{d}f = f_s \mathrm{d}S \tag{10.19}$$

这相当于假设(10.18)泰勒级数展开中 ΔS 很小但不可忽略，而含有$(\Delta S)^2$，$(\Delta S)^3$，…的更小项在 $S \to 0$ 时可忽略．最后在极限情况下，项 $f_s \mathrm{d}S$ 保持不变，其他所有项都被舍掉了．得到的就是(全)微分公式(10.19)．

该方法的合理性在于，当 ΔS 变小时，$(\Delta S)^2$，$(\Delta S)^3$，…等项变小的速度更快．如图 10-1给出了下列函数的图像：

$$g_1(\Delta S) = \Delta S \tag{10.20}$$

$$g_2(\Delta S) = [\Delta S]^2 \tag{10.21}$$

注意当 ΔS 越来越小时，函数 $g_2(\Delta S)$ 接近 0 的速度要快于 $g_1(\Delta S)$.

因此，在标准微积分中所有包含 dS 高于 1 次幂的项都可忽略，从全导数中舍去．问题在于我们是否可以在随机微积分中进行相同的操作呢?

这个重要问题的答案是否定的．在随机环境下，时间变量 t 仍然是确定性的，因此可以对时间变量应用与确定性微积分中相同的关于小的标准．任何含有 dt 高于 1 次幂的项都可以忽略．

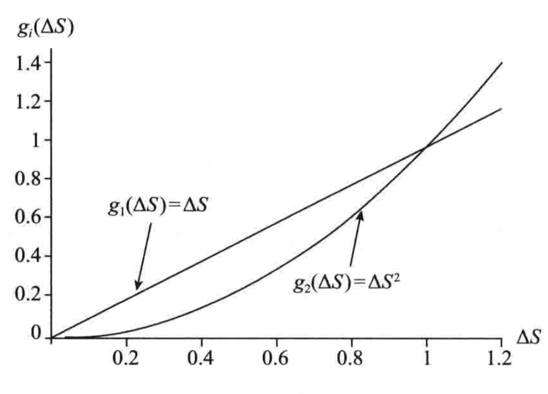


图 10-1　ΔS 与 $(\Delta S)^2$ 的收敛示意图

另一方面，不能对 $\mathrm{d}S_t^2$ 等随机微分使用相同原理．第 9 章已经介绍，在均方意义下，

$$\mathrm{d}W_t^2 = \mathrm{d}t \tag{10.22}$$

因此含有 $\mathrm{d}S_t^2$ 的项可能与 dt 同阶．如果含有 dt 的项保留在泰勒近似式中，随机微分的平方也应该保留．

167

我们对这一点进行进一步的讨论．如果 ΔS_t 是均值为 0 的随机增量，那么 $E[\Delta S_t]^2$ 就是该增量的方差．由于 ΔS_t 是随机的，其方差应该大于 0. 然而该方差表示的是 $(\Delta S_t)^2$ 的"大小". 通常来说，假设 $(\Delta S_t)^2$ 可忽略等价于假设 ΔS_t 的方差趋于 0——这意味着 S_t 几乎是非随机的，这是矛盾的，它使得在市场中对衍生品使用随机微分方程变得没有意义．毕竟，我们的目的是对风险进行定价，而风险源于意料之外的消息．

因此与确定性环境中不同，$(\Delta S_t)^2$ 这类项在随机微分中不可忽略．

假设大小为 h 的量是一阶的，且通常不是小量，那么可以用下面的法则区分可忽略项和不可忽略项．

常用方法：已知函数 $g(\Delta W_k, h)$ 的自变量为维纳过程增量 ΔW_t 和时间增量，考虑比值

$$\frac{g(\Delta W_k, h)}{h} \tag{10.23}$$

如果当 $h \to 0$ 时该比值(在均方意义下)趋于 0，那么我们认为 $g(\Delta W_k, h)$ 在小区间内可忽略；否则，$g(\Delta W_k, h)$ 不可忽略．

该方法相当于将各种项与 h 进行比较．特别地，如果函数 $g(\Delta W_k, h)$ 的均方极限与 h^r $(r>1)$ 成正比，它趋于 0 的速度将快于 h(即小量的平方比小量更小). 另一方面，如果 $r<1$，那么 $g(\Delta W_k, h)$ 的均方极限将与 h 更高次幂而不是 h 成正比．[⊖]

在后面讨论泰勒级数展开时，确定哪些项可以当作小量用到了这个方法．

⊖ 不要忘记函数 $g(\Delta W_k, h)$ 依赖于 ΔW_k 的幂，同时 ΔW_k 的幂也决定了(10.23)中的比值在 h 变小时是否可忽略．这时我们需要讨论泰勒级数展开中的交叉乘积项．

10.3.2 一阶项

再次考虑等式(10.11)：

$$\Delta F(k) = F_s[a_k h + \sigma_k \Delta W_k] + F_t[h] + \frac{1}{2}F_{ss}[a_k h + \sigma_k \Delta W_k]^2 + \frac{1}{2}F_{tt}[h]^2 + F_{st}[h][a_k h + \sigma_k \Delta W_k] \quad (10.24)$$

包含 h 或 ΔS_k 的项显然是不可忽略的一阶增量．用 $F_s[a_k h + \sigma_k \Delta W_k]$或 $F_t h$ 除以 h，当 h 越来越小时，这些项也不消失．例如，

$$\lim_{h\to 0}\frac{F_s a_k h}{h} = F_s a_k \quad (10.25)$$

$$\lim_{h\to 0}\frac{F_t h}{h} = F_t \quad (10.26)$$

显然，它们与 h 无关，不随 h 变小而消失．

另一方面，我们知道比值

$$\frac{F_s \Delta W_k}{h} \quad (10.27)$$

随 h 变小而变大(在概率意义下)，因为 ΔW_k 与 $h^{1/2}$同阶．

因此(10.24)中所有一阶项都不可忽略．

10.3.3 二阶项

用(10.24)右边的二阶项除以 h，考虑比值

$$\frac{F_{tt}h^2}{2h} \quad (10.28)$$

该项仍与 h 成正比，因为在分子中含有依赖于 h^2 的增量——高于一阶的 h 的幂，且该增量是非随机的．因此，该项可忽略： 168

$$\lim_{h\to 0} F_{tt}h = 0 \quad (10.29)$$

接下来考虑含有$[\Delta S_k]^2$ 的二阶项，

$$\frac{1}{2}F_{ss}[\Delta S_k]^2$$

替换 ΔS_k，展开平方项，除以 h，

$$\frac{1}{2}F_{ss}\left[\frac{a_k h^2}{h} + \frac{(\sigma_k \Delta W_k)^2}{h} + \frac{2a_k\sigma_k h\Delta W_k}{h}\right]^2 \quad (10.30)$$

该等式中第一项是小量．分子内含有 h 高于一次的幂，且该项非随机．第三项也是小量．它含有交叉乘积(见下节)．而第二项包含随机变量$(\Delta W_k)^2$．它是均值为 0 的随机变量的平方，在给定过去信息时不可预测，其方差为

$$V[\sigma_k \Delta W_k] = \sigma_k^2 h \quad (10.31)$$

前面写过均方意义下的形式

$$dW_t^2 = dt \tag{10.32}$$

根据定义，我们讨论的 S_k 是随机的，因此 ΔW_k^2 不可被忽略，而 ΔS_k 方差非零意味着

$$\sigma_k > 0 \tag{10.33}$$

最后在使用忽略方法时，

$$\frac{1}{2}F_{ss}\left[\frac{a_k h^2}{h} + \frac{(\sigma_k \Delta W_k)^2}{h} + \frac{2a_k\sigma_k h\Delta W_k}{h}\right]^2 \approx \frac{1}{2}F_{ss}\sigma_k^2 \tag{10.34}$$

该近似式也是在均方意义下的．即在小区间内，等式(10.34)两边差的方差在 $h\to 0$ 时趋于 0.

在得出伊藤公式前，还要对(10.24)泰勒级数展开式中余下的项进行讨论．

10.3.4 含有交叉乘积的项

假设不可预测的部分不包含任何跳跃，那么(10.24)中含有交叉乘积的项在小区间内也是可忽略的，原因在于 S_t 的样本路径的连续性．

考虑将(10.24)中含有交叉乘积的项除以 h：

$$\frac{F_{st}[h][a_k h + \sigma_k \Delta W_k]}{h} = F_{st}[a_k h + \sigma_k \Delta W_k] \tag{10.35}$$

(10.35)右边与 ΔW_k 相关．当 $h\to 0$ 时，ΔW_k 趋于 0. 特别地，由于 $h\to 0$ 时其方差也趋于 0，因此 ΔW_k 可忽略．即 W_k 在 $h=0$ 的极限情况下不发生改变．这是另一种解释维纳过程样本路径连续的方式．

根据前面采用的方法，只要考虑的过程是连续的且不存在跳跃，含有 ΔW_k 和 h 交叉乘积的项就是可忽略的．

10.3.5 余项中的项

余项 R 中的所有项都包含 h 和 ΔW_k 高于 2 次的幂．根据前面的方法，如果不可预测的事件是正常的——即不存在稀有事件——ΔW_k 高于 2 次的幂都是可忽略的．实际上我们在第 8 章中证明过连续鞅和维纳过程的高阶矩在 $h\to 0$ 时可忽略．

169

10.4 伊藤公式

现在，我们总结一下对(10.24)中项的讨论．当 $h\to 0$ 时，舍去全部可忽略项，得到下面的结果：

定理 5 （伊藤引理）$F(S_t,t)$是 t 和随机过程 S_t 的二阶可微函数，其中，S_t 满足

$$dS_t = a_t dt + \sigma_t dW_t$$

a_t 与 σ_t 是具有较好性质的漂移系数与扩散系数．[⊖] 于是有

$$dF_t = \frac{\partial F}{\partial S_t}dS_t + \frac{\partial F}{\partial t}dt + \frac{1}{2}\sigma_t^2\frac{\partial^2 F}{\partial S_t^2}dt \tag{10.36}$$

⊖ 这意味着漂移系数和扩散系数都不会过于不规则，平方可积就满足这一条件．为了书写简便，我们将 $a(S_t,t)$ 记为 a_t，将 $\sigma(S_t,t)$ 记为 σ_t.

其等价的随机微分方程为

$$dF_t = \left[a_t \frac{\partial F}{\partial S_t} + \frac{\partial F}{\partial t} + \frac{1}{2}\sigma_t^2 \frac{\partial F}{\partial S_t}\right]dt + \frac{\partial F}{\partial S_t}\sigma_t dW_t \tag{10.37}$$

该等式在均方意义下成立.

在需要使用伊藤引理的环境中，通常已知随机过程 S_t 的随机微分方程为：

$$dS_t = a(S_t, t)dt + \sigma_t(S_t, t)dW_t \tag{10.38}$$

因此可以将伊藤引理看作已知 S_t 的随机微分方程，求 $F(S_t, t)$ 对应的随机微分方程的工具. 实际上，等式(10.37)是 $F(S_t, t)$ 的随机微分方程.

显然伊藤引理是处理金融衍生品时非常有用的工具. 金融衍生品是建立在标的资产基础上的合约. 只要已知标的资产的随机微分方程，我们就可以利用伊藤引理确定金融衍生品的随机微分方程. 对于市场参与者来说，他们将标的资产价格变化视为外源性的，并希望对金融衍生品进行定价，伊藤公式是必不可少的工具.

10.5 伊藤引理的应用

刚刚提到的伊藤引理的第一种应用，它为得到随机过程函数的随机微分提供了工具.

例如，我们想知道当标的资产价格变化时，期权价格如何变化. 用 $F(S_t, t)$ 表示期权价格，S_t 表示标的资产价格，有

$$dF(S_t, t) = F_s dS_t + F_t dt + \frac{1}{2}F_{ss}\sigma_t^2 dt \tag{10.39}$$

知道 $F(S_t, t)$ 的具体公式，就可以确切地计算出偏导数，并将它们代入上面的公式得到随机微分 $dF(S_t, t)$. 我们在本节的后面部分给出了一些伊藤引理的应用实例.

伊藤引理的第二种应用不太一样，它在计算伊藤积分时很有用. 这一点可能令人觉到有点意外，因为伊藤公式是作为处理随机微分的工具引入的. 在正常情况下，我们并不期望这种公式对于求伊藤积分有什么用处. 在正常积分中积分和导数是分开定义并且服从微积分基本定理的. 但如我们前面指出的那样，随机积分中微分的定义是随机积分的符号化表达. 因此伊藤引理在求随机积分中的实用性并不令人惊讶.

对于伊藤引理的这些应用，我们给出一些简单的例子. 在后面章节中讨论衍生资产定价时，将介绍更多的例子. 170

10.5.1 作为链式法则的伊藤公式

下面讨论的一些简单例子对于熟悉伊藤公式很有帮助.

例 1

考虑标准维纳过程 W_t 的函数

$$F(W_t, t) = W_t^2 \tag{10.40}$$

W_t 的漂移系数为 0，扩散系数为 1.

对该函数应用伊藤公式得

$$dF_t = \frac{1}{2}[2dt] + 2W_t dW_t \tag{10.41}$$

或

$$dF_t = dt + 2W_t dW_t \tag{10.42}$$

注意在这种特殊的情况下，伊藤公式得出的随机微分方程中

$$a(W_t, t) = 1 \tag{10.43}$$

且

$$\sigma(W_t, t) = 2W_t \tag{10.44}$$

因此漂移项是常数，扩散项依赖于信息集 I_t.

例 2

接下来我们对下面的函数应用伊藤公式

$$F(W_t, t) = 3 + t + e^{W_t} \tag{10.45}$$

我们得到

$$dF_t = dt + e^{W_t} dW_t + \frac{1}{2} e^{W_t} dt \tag{10.46}$$

移项得

$$dF_t = \left[\frac{1}{2} e^{W_t} + 1\right] dt + e^{W_t} dW_t \tag{10.47}$$

在该情况下，我们得到 $F(S_t, t)$的随机微分方程，其漂移项、扩散项是与 I_t 相关的：

$$a(W_t, t) = \left[\frac{1}{2} e^{W_t} + 1\right] \tag{10.48}$$

$$\sigma(W_t, t) = e^{W_t} \tag{10.49}$$

10.5.2 作为积分工具的伊藤公式

假设需要计算如下伊藤积分的值，我们在第 9 章中对它进行过讨论：

$$\int_0^t W_s dW_s \tag{10.50}$$

在第 9 章中通过取一些近似和的均方极限直接计算该积分，这种计算方法简单但计算量大. 现在我们利用伊藤引理，通过几步求出该积分.

定义

$$F(W_t, t) = \frac{1}{2} W_t^2 \tag{10.51}$$

对 $F(W_t, t)$应用伊藤公式：

$$dF_t = 0 + W_t dW_t + \frac{1}{2} dt \tag{10.52}$$

它是漂移系数为 1/2、扩散系数为 W_t 的随机微分方程. 对应的积分方程为

$$F(W_t, t) = \int_0^t W_s dW_s + \frac{1}{2} \int_0^2 dt \tag{10.53}$$

求出右边第二个积分项，应用 $F(W_t,t)$ 的定义：

$$\frac{1}{2}W_t^2 = \int_0^t W_s \mathrm{d}W_s + \frac{1}{2}t \tag{10.54}$$

移项，即得到需要的结果

$$\int_0^t W_s \mathrm{d}W_s = \frac{1}{2}W_t^2 - \frac{1}{2}\int_0^t \mathrm{d}t \tag{10.55}$$

171

与第 9 章中利用均方收敛得到的结果相同．

总结一下如何应用伊藤公式计算伊藤积分是非常重要的．

1）猜测函数 $F(W_t,t)$ 的形式．

2）应用伊藤引理得出 $F(W_t,t)$ 的随机微分方程．

3）对新随机微分方程两端同时积分，得到积分方程．[⊖] 该方程中包含的积分较原积分更容易计算．

4）对积分方程移项，得出结果．

该技巧间接但简单．唯一的难点在于得出函数 $F(W_t,t)$ 的确切形式．

我们在后面章节中求积分时都将使用伊藤引理．

另一个例子

假设我们要计算

$$\int_0^t s\,\mathrm{d}W_s \tag{10.56}$$

其中 W_t 仍表示布朗过程．

应用伊藤引理，首先，定义函数 $F(W_t,t)$：

$$F(W_t,t) = tW_t \tag{10.57}$$

对 $F(\cdot)$ 应用伊藤引理，

$$\mathrm{d}F_t = W_t\mathrm{d}t + t\mathrm{d}W_t + 0 \tag{10.58}$$

在对应的积分方程中应用 $\mathrm{d}F_t$ 的定义，

$$\int_0^t \mathrm{d}[sW_s] = \int_0^t W_s\mathrm{d}s + \int_0^t s\,\mathrm{d}W_s \tag{10.59}$$

移项，得到所求的积分：

$$\int_0^t s\,\mathrm{d}W_s = tW_t - \int_0^t W_s\mathrm{d}s \tag{10.60}$$

这里右边第一项来自

$$\int_0^t \mathrm{d}[sW_s] = tW_t - 0 \tag{10.61}$$

再一次，我们通过使用伊藤引理和一系列间接但简单的运算得到了我们所求的积分．

⊖ 实际上，SDE 是积分方程的缩写．因此这一步是要完整地写出具体的 SDE.

10.6 伊藤引理的积分形式

我们反复提到，随机微分是小时间区间上伊藤积分的简写．因此可以将伊藤公式写成积分形式．

对(10.39)两端同时积分，得到

$$F(S_t,t)=F(S_0,0)+\int_0^t\left[F_u+\frac{1}{2}F_{ss}\sigma_u^2\right]\mathrm{d}u+\int_0^t F_s\mathrm{d}S_u \tag{10.62}$$

其中用到等式

$$\int_0^t \mathrm{d}F_u=F(S_t,t)-F(S_0,0)] \tag{10.63}$$

我们可以利用(10.62)版的伊藤公式来得到另一个特性．对(10.62)进行移项，

$$\int_0^t F_s\mathrm{d}S_u=[F(S_t,t)-F(S_0,0)]-\int_0^t\left[F_u+\frac{1}{2}F_{ss}\sigma_u^2\right]\mathrm{d}u \tag{10.64}$$

172
该等式给出了维纳过程或其他连续时间随机过程关于时间的积分函数表达式．要记住在(10.62)和(10.64)中 F_s，F_{ss} 也依赖于 u.

10.7 更复杂环境下的伊藤公式

伊藤公式可以看作一种在给定标的过程 S_t 的随机微分方程的条件下得到函数 $F(S_t,t)$ 随机微分方程的方法．当 $F(S_t,t)$ 为金融衍生品价格、S_t 为标的资产价格时，该工具非常有用．但我们目前所学的内容还不足以胜任金融市场参与者所面对的真实环境．

我们讨论的伊藤公式是建立在单变量的基础上，且假设不可预期的消息可以通过维纳过程增量进行刻画．

可以想象出两种该模型不适用的环境，首先，在某些条件下，函数 $F(\cdot)$ 可能不止依赖于随机变量 S_t，那么就需要用到伊藤公式的多变量版本．这种扩展很简单，但是最好也简单地讨论一下．

第二种不适用的环境概括起来更加复杂，金融市场受稀有事件影响，仅考虑由维纳过程构成的误差项可能并不合适．我们会希望在随机微分方程中计入驱动价格变化跳跃的过程，这种情况下相应的伊藤公式会发生变化，这就是本节所讨论的第二种情况．

10.7.1 多变量情况

现在我们将伊藤公式扩展到多变量结构中，并给出例子．为了简便，我们选取双变量情形，希望读者能够将该公式推广到更多维的情形中．

假设 S_t 是一个服从下面随机微分方程的 2×1 向量随机过程：⊖

$$\begin{bmatrix}\mathrm{d}S_1(t)\\ \mathrm{d}S_2(t)\end{bmatrix}=\begin{bmatrix}a_1(t)\\ a_2(t)\end{bmatrix}\mathrm{d}t+\begin{bmatrix}\sigma_{11}(t) & \sigma_{12}(t)\\ \sigma_{21}(t) & \sigma_{22}(t)\end{bmatrix}\begin{bmatrix}\mathrm{d}W_1(t)\\ \mathrm{d}W_2(t)\end{bmatrix} \tag{10.65}$$

⊖ 类似 $\mathrm{d}S_1\mathrm{d}S_2$ 的项依赖于 $\mathrm{d}W_1(t)\mathrm{d}W_2(t)$.

这意味着我们有如下两个方程：

$$dS_1(t) = a_1(t)dt + [\sigma_{11}(t)dW_1(t) + \sigma_{12}(t)dW_2(t)] \tag{10.66}$$

$$dS_2(t) = a_2(t)dt + [\sigma_{21}(t)dW_1(t) + \sigma_{22}(t)dW_2(t)] \tag{10.67}$$

其中 $a_i(t)$，$\sigma_{ij}(t)(i=1, 2, j=1, 2)$是漂移系数和扩散系数，它们可能依赖于 $S_i(t)$，且 $W_1(t)$，$W_2(t)$是两个相互独立的维纳过程．

在双变量结构中，$S_1(t)$，$S_2(t)$表示受相同维纳成分影响的两个随机过程．由于不同方程的参数 $\sigma_{ij(t)}$ 可能不同，影响两个方程的误差项也会不同．但是由于 $S_1(t)$，$S_2(t)$具有相同的误差成分，除下面的特殊情况外，通常它们是相关的．

$$\sigma_{12}(t) = 0, \quad \sigma_{21}(t) = 0 \text{ 对任意 } t \tag{10.68}$$

假设我们现在有一个连续的、二次可微的函数 $F(S_1(t), S_2(t), t)$，它依赖于 $S_1(t)$，$S_2(t)$. 那么我们如何写出随机微分 dF_t 呢？

多变量形式的伊藤引理给出了答案[㊀]

$$dF_t = F_t dt + F_{s_1} dS_1 + F_{s_2} dS_2 + \frac{1}{2}[F_{s_1 s_1} dS_1^2 + F_{s_2 s_2} dS_2^2 + 2F_{s_1 s_2} dS_1 dS_2] \tag{10.69}$$

173

其中平方微分$[dS_1]^2$，$[dS_2]^2$ 及交叉乘积项 $dS_1 dS_2$ 等于它们的均方极限．

我们已经知道 dt^2 和交叉乘积 $dt\,dW_1(t)$，$dt\,dW_2(t)$在均方意义下等于 0. 这一点在导出单变量伊藤引理时已经讨论过了．这里唯一的不同之处在于交叉乘积 $dW_1(t)dW_2(t)$.[㊁] 我们在这里得到的是两个独立维纳过程增量的乘积．在有限区间 Δ 内，我们希望得到

$$E[\Delta W_1(t)\Delta W_2(t)] = 0 \tag{10.70}$$

因此，可以建立一个极限，使得在均方意义下：

$$dW_1(t)dW_2(t) = 0 \tag{10.71}$$

它为 dS_1^2，dS_2^2 给出了下面的均方近似：

$$dS_1(t)^2 = [\sigma_{11}^2(t) + \sigma_{12}^2(t)]dt \tag{10.72}$$

$$dS_2(t)^2 = [\sigma_{21}^2(t) + \sigma_{22}^2(t)]dt \tag{10.73}$$

交叉乘积项为

$$dS_1(t)dS_2(t) = [\sigma_{11}(t)\sigma_{21}(t) + \sigma_{12}(t)\sigma_{22}(t)]dt \tag{10.74}$$

可以将这些表达式代入(10.69)中的双变量伊藤公式中，消去 dS_1^2，dS_2^2 和 $dS_1 dS_2$.

1. 金融衍生品中的一个例子

标的为债券的期权是最受欢迎的利率衍生品之一．在计算这些衍生品的价格时，收益率曲线起着基础的作用．一种利率期权模型假设收益曲线依赖于两种变量：表示短期利率的 r_t 和表示长期利率的 R_t. 利率衍生品的价格记为 $F(r_t, R_t, t)$，$t \in [0, T]$.

假设这两种利率服从下面的随机微分方程：

$$dr_t = a_1(t)dt + [\sigma_{11}(t)dW_1(t) + \sigma_{12}(t)dW_2(t)] \tag{10.75}$$

㊀ 关于 t 我们对记号做了一点细微的改变．

㊁ 在以下的方程中，我们写随机微分时没有注明它们依赖于 t.

$$dR_t = a_2(t)dt + [\sigma_{21}(t)dW_1(t) + \sigma_{22}(t)dW_2(t)] \tag{10.76}$$

因此短期利率和长期利率的误差项相关．在长度为 h 的有限区间上，其相关系数为

$$\text{Corr}(\Delta r_t, \Delta R_t) = [\sigma_{11}(t)\sigma_{21}(t) + \sigma_{12}(t)\sigma_{22}(t)]h \tag{10.77}$$

市场参与者可以选择参数 $\sigma_{ij}(t)$ 使等式满足观察到的短期和长期利率相关系数和波动性性质.

对这些利率期权定价时，可能需要了解期权价格如何对收益曲线上的微小变化（即 dr_t 和 dR_t）做出反应．也就是说需要得到随机微分 dF_t. 这里要用到多变量形式的伊藤公式[⊖]：

$$\begin{aligned} dF_t &= F_t dt + F_r dr + F_R dR + \frac{1}{2}[F_{rr}(\sigma_{11}^2 + \sigma_{12}^2) \\ &\quad + F_{RR}(\sigma_{21}^2 + \sigma_{22}^2) + 2F_{rR}(\sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22})]dt \end{aligned} \tag{10.78}$$

给定收益率曲线上的一个由 dr_t 和 dR_t 确定的微小变化，随机微分 dF_t 可以度量利率衍生品价格如何在小区间 dt 上发生变化．

2. 财富

174

投资者以价格 $P_i(t)$购买 $N_i(t)$单位第 i 种资产．共有 n 种资产，$N_i(t)$和 $P_i(t)$都是连续时间的随机过程，它们可能是自变量相同的函数．

投资总价值由 t 时刻财富函数 $Y(t)$给出：

$$Y(t) = \sum_{i=1}^{n} N_i(t)P_i(t) \tag{10.79}$$

假设我们要计算财富随时间的增长量．应用伊藤引理：

$$dY(t) = \sum_{i=1}^{n} N_i(t)dP_i(t) + \sum_{i=1}^{n} dN_i(t)P_i(t) + \sum_{i=1}^{n} dN_i(t)dP_i(t) \tag{10.80}$$

显然，如果使用标准微积分公式，等式最后一项是不存在的．

10.7.2 伊藤公式和跳跃

前面我们总是假定标的过程 S_t 可以表示为维纳过程的函数，该假设可能过于严格．因为随机误差项中也可能存在跳跃成分，我们将在本节中给出伊藤公式的推广形式．

假设我们观察随机过程 S_t，它服从下面的随机微分方程：

$$dS_t = a_t dt + \sigma_t dW_t + dJ_t, \quad t \geqslant 0 \tag{10.81}$$

其中 dW_t 为标准维纳过程．新出现的项 dJ_t 表示可能存在的非预期跳跃．该跳跃部分在有限区间 h 上均值为 0：

$$E[\Delta J_t] = 0 \tag{10.82}$$

我们进行这样的假设是因为该项是不可预测的干扰项的一部分．由于跳跃中任何可预测部分都被包含在漂移项 a_t 中，因此该假设是非限制性的．

⊖ 为了书写简便，我们还将 $\sigma_{ij}(t)$写作 σ_{ij}.

假设跳跃具有下面的结构：J_t 在两个跳跃之间为常数．在跳跃时刻 $\tau_j(j=1, 2, \cdots)$ 为随机离散的．只要发生跳跃，跳跃是随机发生的且是相互独立的，幅度为 a_i 的跳跃发生的概率为 p_i．[㊀]

因此在微小有限区间 h 内，增量 ΔJ_t（近似地）为

$$\Delta J_t = \Delta N_t - \left[\lambda_t h\left(\sum_{i=1}^{k} a_i p_i\right)\right] \tag{10.83}$$

其中 N_t 表示到时刻 t 为止所有跳跃的幅度和，它是一个随机过程．具体而言，如果区间 h 内只有一个跳跃且值为 a_i，那么 N_t 的值为 a_i．项 $\sum_{i=1}^{k} a_i p_i$ 为跳跃的期望值，笼统地说，$\lambda_t h$ 表示发生跳跃的概率．ΔN_t 减去它们是为了使 ΔJ_t 变为不可预测．

在此条件下，可以认为漂移系数 t 表示两种独立的漂移和，一种属于维纳连续成分，另一种是 S_t 发生的纯跳跃

$$a_t = \alpha_t + \lambda_t\left(\sum_{i=1}^{k} a_i p_i\right) \tag{10.84}$$

其中 α_t 是 S_t 中连续运动的漂移系数．

跳跃过程必须再次进行讨论．跳跃的发生是随机的，但一旦发生跳跃，其幅度也是随机的．此外，上面给出的结构假设两种随机过程相互独立．

175

在此条件下，伊藤公式为

$$\mathrm{d}F(S_t,t) = \left[F_t + \lambda_t \sum_{i=1}^{k}(F(S_t + a_i, t) - F(S_t,t))p_i + \frac{1}{2}F_{ss}\sigma^2\right]\mathrm{d}t + F_s\mathrm{d}S_t + \mathrm{d}J_F \tag{10.85}$$

其中 $\mathrm{d}J_F$ 为

$$\mathrm{d}J_F = [F(S_t,t) - F(S_t^-,t)] - \lambda_t\left[\sum_{i=1}^{k}(F(S_t + a_i, t) - F(S_t,t))p_i\right]\mathrm{d}t \tag{10.86}$$

定义 S_t^- 为

$$S_t^- = \lim_{s\to t} S_s, \quad s < t \tag{10.87}$$

它表示 S 在 t 之前无穷小时刻的值．

实务中如何计算 $\mathrm{d}J_F$ 呢？首先要计算等式(10.86)中右边第二项，有可能发生的随机跳跃引起的变化期望值．为此，要同时使用 $\mathrm{d}t$ 期间发生跳跃的概率和由 S_t 引起的 $F(\cdot)$ 跳跃的期望值．若在某个特定时间段内观察到跳跃，则右边第一项也要被包括进去．否则该项为 0.

10.7.3 半鞅的伊藤引理

半鞅是局部鞅加有界变差过程(Protter，1990)．特别地，有界变差过程本身就是一个半

㊀ 在标准泊松过程中，所有跳跃的幅度都为 1，这一步是多余的．

鞅. 众所周知，扩散过程是半鞅，若 X 是半鞅，$f\in C^2$，则$(f(X_t))_{t\geqslant 0}$也是半鞅，且

$$f(X_t)-f(X_0)=\int_0^t f'(X_s)\mathrm{d}X_s+\frac{1}{2}\int_0^t f''(X_s)\mathrm{d}X_s + \sum_{0<s\leqslant t}(f(X_s)-f(X_{s-}-f'(X_{s-}))\Delta X_s) \tag{10.88}$$

令 $X=(X_1,\cdots,X_n)$为 n 维连续半鞅，$f:R^n\to R$ 具有连续二阶偏导数．那么 $f(X)$也为半鞅，且有下列分解公式：

$$f(X_t)-f(X_0)=\sum_{i=1}^n\int_0^t\frac{\partial f}{\partial x_i}(X_s)\mathrm{d}X_s^i+\frac{1}{2}\sum_{1\leqslant i,j\leqslant n}\int_0^t\frac{\partial^2 f}{\partial x_i\,\partial x_j}\mathrm{d}[X^i,X^j]_s \tag{10.89}$$

其中$[X^i,X^j]$是 X^i 和 X^j 的二次变差．

这种伊藤公式在 Protter(1990，p. 74)中提到过．此外，Protter(1990，p. 68)证明了一个定理，如果 X 是连续有界变差过程，那么对于任意半鞅 Y，X 和 Y 的二次变差$[X,Y]$对于任意 t 都等于 X_0Y_0. 因此 $\mathrm{d}[W,M]_t=0$，$\mathrm{d}[M,M]_t=0$.

10.8 小结

伊藤引理是随机微积分中的核心微分工具，一些基础知识要记住．首先，已知标的资产变化时，该公式有助于确定金融衍生品的随机微分．其次，该公式完全依赖于伊藤积分的定义，这意味着应该在随机等价的环境下对等式进行解释．

176

最后从实际应用的角度看，读者们应该记住，确定性微积分中标准公式得出的结果与伊藤公式有显著的不同．特别地，使用标准公式相当于假设所有观察过程连极小的波动性都没有．当我们试图为含风险的金融衍生品定价时，这并不是一个令人愉快的假设．

10.9 参考阅读

第 9 章中推荐的资料同样适用于本章．伊藤引理和伊藤积分是两个常被同时考虑的话题. 读者还可以参考 Kushner 和 Dupuis(2000)，它给出了带跳跃过程的伊藤引理的例子．

10.10 习题

1. 计算下列函数关于维纳过程 W_t 的微分，如果存在，计算出关于 t 的微分．

(a) $f(W_t)=W_t^2$

(b) $f(W_t)=\sqrt{W_t}$

(c) $f(W_t)=\mathrm{e}^{(W_t^2)}$

(d) $f(W_t,t)=\mathrm{e}^{\left(\sigma W_t-\frac{1}{2}\sigma^2 t\right)}$

(e) $f(W_t)=\mathrm{e}^{\sigma W_t}$

(f) $g=\int_0^t W_s\,\mathrm{d}s$

2. 假设 $W_{t_i}(i=1,2)$是两个维纳过程．利用伊藤引理为下列变换得出合适的随机微分方程．

(a) $X_t=(W_{t_1})^4$

(b) $X_t=(W_{t_1}+W_{t_2})^2$

(c) $X_t=t^2+\mathrm{e}^{W_{t_2}}$

(d) $X_t=\mathrm{e}^{t^2+W_{t_2}}$

3. W_t 为维纳过程．再次考虑几何过程 S_t：

$$S_t = S_0 \mathrm{e}^{\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t}$$

(a) 计算 $\mathrm{d}S_t$.

(b) S_t 的期望变化率是多少？

(c) 如果 S_t 定义中的指数项中没有包含 W_t 项，$\mathrm{d}S_t$ 将为多少？S_t 的期望变化又将为多少？

4. W 为维纳过程，计算 $E(W_t^4)$.

5. 考虑等式 $X_t = \cos(\mathrm{e}^{W_t})$. 编写一个程序，通过拟合计算估计 X_t 的均值和方差. 177

第 11 章　衍生品价格的动态变化

11.1　引言

我们在第 7 章介绍了随机微分方程的概念．随后在第 9 章中我们通过伊藤积分把这一概念(在数学意义上)进行了正规化．从而，我们得到表达式

$$dS_t = a(S_t,t)dt + \sigma(S_t,t)dW_t, \quad t \in [0,\infty) \tag{11.1}$$

来作为

$$\int_t^{t+h} dS_u = \int_t^{t+h} a(S_u,u)du + \int_t^{t+h} \sigma(S_u,u)dW_u \tag{11.2}$$

当 h 趋向于无穷小时的符号化表达．

我们需要再次强调在这个推导过程中的几个值得注意的问题．首先，在得到(11.1)的过程中，我们并没有使用有关任何金融市场的理论和概念．我们的基本工具只是伊藤积分，所用到的方法是将资产价格的随机增量分为两个部分：可预测的增量和不可预测的增量．

179

这就引出了第二点．等式(11.1)的分解使用了 t 时刻已知的信息集．由于不同的交易者可能拥有不同的信息集，那么不同的交易者所能够推算出的形式为(11.1)的随机微分方程也可能是不同的．例如，考虑下面这个极端的情况．假设某个市场参与者拥有内幕消息，能够事先获悉所有会引起价格变化的随机事件．在这个(不现实的)条件下，(11.1)中的扩散项将为 0. 这是因为他知道 dS_t 将如何变化，就能准确预测该变量，从而对于所有 t，$dW_t=0$. 因此，对于这位市场参与者而言，资产价格所服从的随机微分方程为

$$dS_t = a^*(S_t,t)dt \tag{11.3}$$

而对于所有其他市场参与者而言，相应的随机微分方程仍然是

$$dS_t = a(S_t,t)dt + \sigma(S_t,t)dW_t \tag{11.4}$$

这两个方程的漂移项、扩散项不可能相同．这是因为，作为在随机微分方程中起主要作用的那一部分，误差项的不同决定了 $a^*(S_t,t)$ 与 $a(S_t,t)$ 也不相同．这个例子显示了在随机微分方程的表达式中误差项 dW_t 的定义总是依赖于信息集族$\{I_t, t\in[0,\infty)\}$的．如果所能够得到的信息集族$\{I_t^*\}$不同，我们对于资产价格的预测误差也将会是不同的，那么我们得到的随机微分方程(1.11)中的误差项所表达的随机行为也应该产生相应的变化．所以，当我们得到一个新的信息集族时，就要用一个新的误差项 dW_t^* 来代替 dW_t，而 dW_t^* 可能具有比 dW_t 更小的方差．

在随机微积分中，我们可以把 W_t 的这种性质严格地表达为：维纳过程 W_t 是和信息集族 I_t 相适应的．

由于我们可以通过随机微分方程构造规范化的(数学)模型来模拟标的资产随时间所产

生的价格变化，因此在其衍生资产的定价中也经常会用到随机微分方程．此外，我们推导随机微分方程的过程也与金融市场中实际交易者的行为是一致的：在一个交易日中，交易者会不断尝试预测资产价格，并且记录发生的新事件；而这些事件中总是会包含一些我们在观察到 dS_t 之前无法预测的部分．而在观察到之后，它们就变为已知，并且成为交易者的新信息集的一部分．

本章将介绍随机微分方程的一些性质．

关于 a_t 和 σ_t 的条件

随机微分方程

$$dS_t = a(S_t,t)dt + \sigma(S_t,t)dW_t, \quad t\in[0,\infty) \tag{11.5}$$

的漂移项和扩散项可以与 S_t 和 t 有关，因此，这些参数本身就是随机变量．值得注意的是，如果已知 t 时刻的信息，这些参数可以由市场参与者的观察得到，那么此时它们就是常量，而非随机变量．这也是"参数是和 I_t 相适应的"这一重要假设的结果．此外，如同之前章节中已多次提到的，我们总是假设这些参数是良态的．

按照惯例，每次构建随机微分方程模型时，我们都要具体给出这些正则性条件．

我们总是假设 $a(S_t,t)$和 $\sigma(S_t,t)$满足以下条件：

$$P\left(\int_0^t |a(S_u,u)|\,du < \infty\right) = 1 \quad 且 \quad P\left(\int_0^t |\sigma(S_u,u)|\,du < \infty\right) = 1$$

这些条件有相似的含义．它们要求漂移和扩散参数不随时间产生"太多"变化．

这些条件中的积分只是关于时间变量的积分，在通常情况下，它们都是可定义的．此时这些条件意味着漂移和扩散参数是概率为 1 的有界变差函数． 180

注意，在本书的剩余部分，我们总是默认这些条件成立，且不再重复说明．

11.2 随机微分方程对应路径的几何描述

考虑随机微分方程

$$dS_t = a(S_t,t)dt + \sigma(S_t,t)dW_t, \quad t\in[0,\infty) \tag{11.6}$$

其中漂移和扩散参数依赖于观察到的资产价格 S_t(可能也同时依赖于时间变量 t).

这种随机微分方程中 S_t 对应的几何图像是怎样的?

图 11-1 描述了一个例子．我们考虑长度为 h 的离散小区间．可以看到，S_t 随时间的运动可以分解为两部分．第一，时间点之间有一条可预测的路径，由带箭头的斜线标出．然后，在每个 $t_k = kh$，有第二种运动，

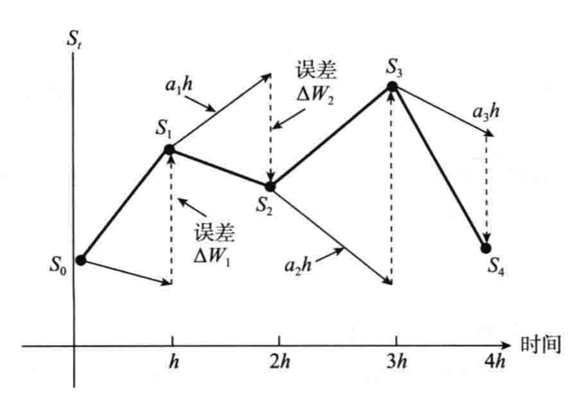


图 11-1 几何随机微分方程在长度为 h 的离散区间上的样本路径

且与第一种运动垂直,[⊖]由竖直方向的箭头表示．这两种运动有时方向相反，有时方向相同．S_t 的真实运动由这两部分共同组成，在图中用粗线表示．

该几何推导再次说明当 h 无穷小时，S_t 的变化可能会非常复杂．

11.3 随机微分方程的求解

随机微分方程归根结底是一个*方程*．也就是说，它含有未知量，就是随机过程 S_t．随机微分方程的求解比看起来的复杂得多．我们所求的并不是一个数或者一个向量，而是一个随机过程，它的轨迹及其相关联的概率都需要准确确定．

11.3.1 解意味着什么

首先，我们考虑用小区间上的离散差分来近似原来的随机微分方程：

$$S_k - S_{k-1} = a(S_{k-1}, k)h + \sigma(S_{k-1}, k)\Delta W_k, \quad k = 1, 2, \cdots, n \tag{11.7}$$

该方程的解是随机过程 S_t．我们希望找到一列以 k 为下标的随机变量 S_k，使增量 ΔS_k 满足(11.7)．此外，我们还想知道满足(11.7)的 S_k 的矩和分布函数．从表面上看，已知某个 $a(\cdot)$和 $\sigma(\cdot)$，我们不知道能否求出一列随机数，使其轨迹对任意 k 都满足(11.7)．

更重要的是，我们的最终目标是求 h 趋于 0 时的解．如果一个连续时间过程 S_t 对任意 $t>0$都满足方程

$$\int_0^t dS_u = \int_0^t a(S_u, u)du + \int_0^t \sigma(S_u, u)dW_u \tag{11.8}$$

那么我们说 S_t 是

$$dS_t = a(S_t, t)dt + \sigma(S_t, t)dW_t \tag{11.9}$$

的解．由于随机微分方程的解是随机过程，求解随机微分方程与求解普通的微分方程本质上有很大差别．实际上，在随机微积分中，有两类不同类型的解的定义．

11.3.2 解的种类

181

随机微分方程的第一类解与通常意义下常微分方程(ODE)的解类似：已知漂移、扩散参数和随机干扰项 dW_t，我们需要确定随机过程 S_t，使其满足随机微分方程：

$$dS_t = a(S_t, t)dt + \sigma(S_t, t)dW_t, \quad t \in [0, \infty) \tag{11.10}$$

显然，这样的解 S_t 依赖于时间 t 以及随机变量 W_t 的过去和当前值．对 $t>0$，它所对应的积分表达式为

$$\int_0^t dS_u = \int_0^t a(S_u, u)du + \int_0^t \sigma(S_u, u)dW_u \tag{11.11}$$

这类解给出了这种关系的具体表达式．当(11.11)右端的 W_t 为外生，继而确定 S_t 时，我们称得到了该随机微分方程的*强解*．这与常微分方程的解很相似．

⊖ 这里“垂直”指的是概率中所说的“不相关的”．

第二类解则是随机微分方程所特有的，它叫做*弱解*．在弱解中，我们可以得到过程 $\widetilde{S}_t$：

$$\widetilde{S}_t = f(t, \widetilde{W}_t) \tag{11.12}$$

其中 $\widetilde{W}_t$ 是维纳过程，其分布与 $\widetilde{S}_t$ 同步确定．因此，在考虑随机微分方程的弱解时，我们所拥有的已知条件只有漂移和扩散参数 $a(\cdot)$和 $\sigma(\cdot)$．

弱解的思想可以这样理解：既然求解随机微分方程必须找到满足(11.8)的随机变量，我们可以考虑同时寻找 $\widetilde{S}_t$ 和 $\widetilde{W}_t$，使得一组$\{\widetilde{S}_t, \widetilde{W}_t\}$满足方程，这也是一类随机微分方程的解．

因此，所谓求弱解是指，我们在已知漂移参数 $a(S_t,t)$和扩散参数 $\sigma(S_t,t)$的情况下，去寻找满足方程(11.8)的 $\widetilde{S}_t$ 和 $\widetilde{W}_t$．而求强解时，W_t 也是一个已知项，不需要我们去求解．

这里还有一些可能让人困惑的问题．首先，既然 dW_t 和 $d\widetilde{W}_t$ 都是具有零均值和方差 dt 的维纳过程，它们有什么区别？它们不是相同的吗？

182

如果考虑它们的分布函数，我们可以很容易地回答这个问题．由于 dW_t 和 $d\widetilde{W}_t$ 的分布函数是一样的，从这个角度上来讲，这两个随机误差没有任何区别．它们之间的差别只在于用来定义 dW_t 和 $d\widetilde{W}_t$ 的信息集．[㊀] 尽管对应的密度函数可能相同，但是只要所适应的信息集不同，这两个随机过程可能表达的就是完全不同的现实现象．

我们需要更清楚地来说明这一点．因为只有明确了这一点，我们才能够更好地理解连续时间下的随机模型．考虑以下随机微分方程，其中扩散项包括直接给出的 dW_t：

$$dS_t = a(S_t,t)dt + \sigma(S_t,t)dW_t \tag{11.13}$$

可以这么说，误差过程 dW_t 描述了那些完全不可预测却又能影响价格的微小事件．而由这些微小事件累计所形成的"历史"就构成了我们在 t 时刻拥有的信息集，我们将其记作 I_t．[㊁]

强解是在给定 dW_t 的情况下求解满足(11.13)的 S_t．也就是说，为得到强解 S_t，我们首先需要知道 I_t 族．这就意味着强解 S_t 必须是 I_t 适应的．

另一方面，弱解的计算并不需要用到某个生成信息集 I_t 的过程，它是伴随着某个 $\widetilde{W}_t$ 得到的．而这个误差过程可能会生成另一个信息集 H_t，因此，弱解不一定是 I_t 适应的，但它关于信息集 H_t 依然是鞅．[㊂]

所以，弱解满足

$$dS_t = a(S_t,t)dt + \sigma(S_t,t)d\widetilde{W}_t \tag{11.14}$$

其中漂移和扩散参数与(11.13)相同，且与某个信息集 H_t 相适应．

11.3.3 哪一种解更好

因为强解和弱解有着相同的漂移和扩散参数，所以 S_t 和 $\widetilde{S}_t$ 有着相似的统计性质．只

㊀ 我们将会在第 14 章看到，两个不同的维纳过程可能对应两个不同的关于 dS_t 的概率测度．

㊁ 如之前所说，在数学中这种信息集叫做 σ-域或者 σ-代数．

㊂ 由于这种鞅性质，SDE 下的伊藤积分仍然能按相同方式定义．

知道均值和方差，我们还不能区分这两类解．但这两类解确实是不同的．[⊖]

使用强解意味着误差项 W_t 是已知的．如果该条件满足，金融分析师更可能选择研究强解.

但在更一般的情况下，当我们需要通过求解随机微分方程来对衍生品定价时，我们实际上并不知道确切的 W_t. 我们所知道的可能只有波动率和(有时候)漂移参数．因此，在这样的条件下进行衍生品定价时，我们研究弱解．

11.3.4 关于强解的讨论

我们之前说过，随机微分方程归根结底是一个方程．这意味着它包含待求的未知量．在随机微分方程中，待求的未知量是随机过程．通过求随机微分方程，我们得到 S_t 对于所有 t 都满足积分方程

$$S_t = S_0 + \int_0^t a(S_u, u)\mathrm{d}u + \int_0^t \sigma(S_u, u)\mathrm{d}W_u \tag{11.15}$$

183 也就是说，假定初始价格是 S_0，资产价格的变化过程是由方程右边的两个积分项决定的．如果得到的解为 S_t，那么 $S_t - S_0$ 必须等于这两个积分项之和．这是检验解的方式．

这个方法用积分方程的形式来检验解的正确性，而不是直接用随机微分方程来检验．为什么这样做呢？事实上，截止到目前的讨论，我们并没有在随机环境下微分的理论．因此，如果有一个随机微分方程的候选解，我们不能直接对其求导来看它是否满足随机微分方程．这个时候我们有两种选择．

我们可以通过一个确定性方程的例子来理解检验随机微分方程的解的过程．我们首先考虑一个简单的常微分方程

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = aX_t \tag{11.16}$$

其中 a 为常数、X_0 已知．没有随机干扰项，不是随机微分方程，我们可以直接检验一个函数是否是方程的解．例如，假如我们怀疑(11.17)

$$X_t = X_0 \mathrm{e}^{at} \tag{11.17}$$

是否是(11.16)的一个解，那么该解必须满足两个条件．第一，如果我们对 X_t 关于 t 求导，导数必须等于 X_t 的 a 倍．第二，如果令 $t=0$，X_t 应该等于 X_0，即已知的初始点．

那么我们来检验它是否是方程(11.16)的解．直接对 X_t 求导

$$\frac{\mathrm{d}}{\mathrm{d}t}(X_0 \mathrm{e}^{at}) = a[X_0 \mathrm{e}^{at}] \tag{11.18}$$

其导数确实是 X_t 的 a 倍，第一个条件满足．

令 $t=0$，我们有

$$(X_0 \mathrm{e}^{a0}) = X_0 \tag{11.19}$$

⊖ 强解同时也是一种弱解，但反过来不成立．

所以候选解也满足初始条件．所以我们可以说 X_t 是常微分方程(11.16)的解．这个方法就是用求导法检验解．㊀

如果没有关于连续随机过程的微分理论，这样一套检验解的方法是无法运用到随机微分方程中的．事实上，如果我们使用同样的微分方法，(错误地)认为它在随机环境下仍然成立，并通过求导来验证随机微分方程的解，那我们将会得到错误的答案．如之前所说，确定性函数中的微分法则在随机变量函数中不成立．

我们接下来再作一些有用的介绍．X_0 已知的常微分方程

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = aX_t \tag{11.20}$$

方程两边都含有未知函数 X_t. 正是如此，ODE 才是一个方程．这个方程的解是一个具体的函数，它依赖于 ODE 中的其余参数和已知变量．参数是$\{a, X_0\}$，而唯一的已知变量是时间 t. 所以未知函数 X_t 的解包括了已知量．其中一个解就是

$$X_t = X_0 \mathrm{e}^{at} \tag{11.21}$$

检验该解需要对右边关于 t 求导，然后看是否满足 ODE.

现考虑以下随机微分方程特例：

$$\mathrm{d}S_t = a\mathrm{d}t + \sigma \mathrm{d}W_t, \quad t \geqslant 0 \tag{11.22}$$

184

其中 S_0 已知．㊁如果已求出该随机微分方程的强解，它将是时间 t、参数$\{a, \sigma, S_0\}$和 W_t 的函数 $f(\cdot)$：

$$S_t = f(a, \sigma, S_0, t, W_t) \tag{11.23}$$

这个解依赖于随机过程 W_t，因此它也是一个随机过程：㊂

如果我们想用确定性微分公式来检验它是否满足(11.22)中的随机微分方程，就需要对 S_t 和 W_t 关于 t 求导．但这些关于 t 的导数并没有明确的定义．所以我们不能照搬确定性问题中的方法来检验解．

取而代之的是，我们要对候选解使用伊藤引理，来看这个候补解是否满足随机微分方程所对应的积分方程．在下面的例子中，我们将对这一点作具体研究．

11.3.5 随机微分方程解的检验

我们还是考虑以下这个随机微分方程

$$\mathrm{d}S_t = \mu S_t \mathrm{d}t + \sigma S_t \mathrm{d}W_t \tag{11.24}$$

Black-Scholes(1973)曾在看涨期权定价中使用该方程．这里 S_t 表示某无红利证券的价格．

两边同时除以 S_t，得

㊀ 当然，读者一定会好奇候选解是怎么来的．这个问题归属于微分方程．这里，我们仅讨论金融中使用的模型．

㊁ 漂移和扩散参数是常数，不依赖于 t 时刻的信息．

㊂ 要记住以上讨论的 SDE 是一个特例．通常，像(11.23)那样的 S_t 的强解，一般依赖于 $a(S_u, u)$、$\sigma(S_u, u)$和 $\mathrm{d}W_u$ 的积分．这种依赖关系在整个 W_t 的轨迹都存在．

$$\frac{1}{S_t}\mathrm{d}S_t = \mu \mathrm{d}t + \sigma \mathrm{d}W_t \tag{11.25}$$

我们首先来计算其所对应的积分方程：

$$\int_0^t \frac{\mathrm{d}S_u}{S_u} = \int_0^t \mu \mathrm{d}u + \int_0^t \sigma \mathrm{d}W_u \tag{11.26}$$

由于右端第一个积分不包含随机项，可以用常规方法计算：

$$\int_0^t \mu \mathrm{d}u = \mu t \tag{11.27}$$

第二个积分包含随机项，但 $\mathrm{d}W_t$ 的系数是不随时间变化的常数，所以其积分也可以按照一般方法得到

$$\int_0^t \sigma \mathrm{d}W_u = \sigma[W_t - W_0] \tag{11.28}$$

其中根据定义，$W_0=0$，所以有

$$\int_0^t \frac{\mathrm{d}S_u}{S_u} = \mu t + \sigma W_t \tag{11.29}$$

任何原随机微分方程的解都必须满足这个积分方程．在本例中，我们能用伊藤引理轻松地证明这一点．

考虑候选解

$$S_t = S_0 \mathrm{e}^{\{(a-\frac{1}{2}\sigma^2)t+\sigma W_t\}} \tag{11.30}$$

注意这个候选解是含有参数 a、时间 t 和随机变量 W_t 的函数．由于 S_t 依赖于 I_t 且是 I_t 适应的，我们所研究的显然是一个强解．

那么我们应该如何验证该函数确实是一个解呢？

我们用伊藤引理可以计算随机微分 $\mathrm{d}S_t$：

$$\mathrm{d}S_t = \left[S_0 \mathrm{e}^{\{(a-\frac{1}{2}\sigma^2)t+\sigma W_t\}}\right]\left[\left(a-\frac{1}{2}\sigma^2\right)\mathrm{d}t + \sigma \mathrm{d}W_t + \frac{1}{2}\sigma^2 \mathrm{d}t\right] \tag{11.31}$$

185 其中右端最后一项所对应的是伊藤引理中的二阶项．

化简这个表达式并将 S_t 代回，我们可以得到

$$\mathrm{d}S_t = S_t[a\mathrm{d}t + \sigma \mathrm{d}W_t] \tag{11.32}$$

若 $a=\mu$，它就是原随机微分方程．有趣的是，在使用伊藤引理的过程中，$\frac{1}{2}\sigma^2$ 项被消掉了．如果我们使用的是确定性微分中的规则，该项在(11.32)中不会消失，那么(11.30)中的函数不满足随机微分方程．

事实上，如果我们使用常规微积分，得到的全微分为

$$\mathrm{d}S_t = S_t\left[\left(a-\frac{1}{2}\sigma^2\right)\mathrm{d}t + \sigma \mathrm{d}W_t\right] \tag{11.33}$$

若 $a=\mu$，(11.33)与原随机微分方程不同．所以在使用常规微积分的情况下，我们会错误地认为函数(11.30)不是随机微分方程(11.24)的解．

11.3.6 一个重要的例子

假如 S_t 是某个资产的价格过程，它具有随机升值率，也就是说，我们有

$$dS_t = rS_t dt + \sigma S_t dW_t, \quad t \in [0, \infty) \tag{11.34}$$

上一节我们讨论了这个随机微分方程的一个候选(强)解

$$S_t = S_0 e^{\{(r-\frac{1}{2}\sigma^2)t+\sigma W_t\}} \tag{11.35}$$

现在，假设 S_T 是该资产在某个未来时刻 $T>t$ 的价格．那么，在时刻 t，S_T 是一个未知量．但是它是可以被预测的，而最好的预测方法就是使用条件概率：

$$E_t[S_T] = E[S_T | I_t] \tag{11.36}$$

在资产定价理论中，人们关心以下等式是否成立

$$S_t = e^{-r(T-t)} E_t[S_T] \tag{11.37}$$

这个等式意味着资产在未来时刻 T 期望价格的折现率为 r 的折现值应该等于资产现价．这个鞅性质很重要，因为它可以用于计算现价 S_t.

现在我们计算 $E_t[S_T]$. 首先我们可以把 S_T 分解为

$$S_T = [S_0 e^{(r-\frac{1}{2}\sigma^2)T}][e^{\sigma W_T}] \tag{11.38}$$

这样一来，S_T 的期望值只依赖于

$$e^{\sigma WT} \tag{11.39}$$

的期望值．注意，这个表达式是一个关于 W_T 非线性的函数．所以，S_T 也是关于 W_T 非线性的．因此，在求解期望 $E_t[S_T]$时，我们不能将期望运算 $E_t[\cdot]$直接“转移”到随机项 W_T 上面.

有用两种不同的方法来求这个期望值．一种方法是利用维纳过程 W_T 的密度函数，然后直接用积分求期望

$$E_t[e^{\sigma W_t}] = \int_{-\infty}^{\infty} e^{\sigma W_T} f(W_T | W_t) dW_T \tag{11.40}$$

其中积分中的 $f(W_T | W_t)$是 W_T 的(条件)密度函数．其(条件)均值是 W_t，方差是 $T-t$．

计算该积分并不困难．但我们更偏向第二种方法，这是一种可以体现“随机微积分”特色的计算方法．我们将会再次用到伊藤引理，并引入一个随机微积分中常用的积分方程．

根据(11.38)，S_T 可表示为

$$S_T = [S_0 e^{(r-\frac{1}{2}\sigma^2)T}][e^{\sigma W_T}] \tag{11.41}$$

第二种计算方法的基本思想是将关于 W_t 的非线性函数转化为关于 W_t 的线性函数，然后我们就直接求期望，而不需要对维纳过程的密度函数做积分运算．

186

虽然这个方法是间接的，但是使用起来却相当方便．首先将(11.38)中的非线性随机项记为 Z_t：[⊖]

⊖ 在这之后的运算中，我们(不失一般性地)用下标 t 代替 T. 这样做可以简化表达式．

$$Z_t = e^{\sigma W_t} \tag{11.42}$$

然后对 Z_t 使用伊藤引理，得

$$dZ_t = \sigma e^{\sigma W_t} dW_t + \frac{1}{2}\sigma^2 e^{\sigma W_t} dt \tag{11.43}$$

接下来，运用对应的积分方程：

$$Z_t = Z_0 + \sigma\int_0^t e^{\sigma W_s} dW_s + \int \frac{1}{2}\sigma^2 e^{\sigma W_s} ds \tag{11.44}$$

最后，对两边作期望，且注意：

$$E[Z_0] = 1 \tag{11.45}$$

这是由于 $W_0=0$ 和 Z_t 的定义．此外，因为维纳过程的增量与过去的观察值无关，我们有

$$E\left[\int_0^t e^{\sigma W_s} dW_s\right] = 0 \tag{11.46}$$

所以，我们可以得到

$$E[Z_t] = 1 + \int \frac{1}{2}\sigma^2 E[Z_t] ds \tag{11.47}$$

这里我们根据 Z_t 的定义对(11.44)的第二项积分做了 $e^{\sigma W_s}=Z_s$ 的变量代换．

方程(11.47)有一些有意思的特点．首先，这个方程不包含关于随机变量的积分．其次，这个方程关于 $E[Z_t]$是线性的．所以它能够用求解一般常微分方程的方法来求解．例如，如果把 $E[Z_t]$看作一个确定性变量，记作为 X_t，那么有

$$x_t = 1 + \int_0^t \frac{1}{2}\sigma^2 x_s ds \tag{11.48}$$

这个积分方程等价于常微分方程[⊖]

$$\frac{dx_t}{dt} = \frac{1}{2}\sigma^2 x_t \tag{11.49}$$

其中初始条件 $x_0=1$. 易知该常微分方程的解是

$$x_t = E[Z_t] = e^{\frac{1}{2}\sigma^2 t} \tag{11.50}$$

其中 $E[Z_0]=1$. 回到 $E_t[S_T]$，有

$$E_t[S_T] = [S_0 e^{(r-\frac{1}{2}\sigma^2)T}]E[Z_t] \tag{11.51}$$

利用我们刚刚得到的关于 $E[Z_t]$的结果，可得

$$E_t[S_T] = [S_0 e^{(r-\frac{1}{2}\sigma^2)T}]e^{\sigma W_t} e^{\frac{1}{2}\sigma^2(T-t)} \tag{11.52}$$

其中方程右边出现关于 W_t 的那一项是因为这里的取值是以 t 时刻的信息为条件．由于

$$S_T = S_0 e^{(r-\frac{1}{2}\sigma^2)t+\sigma W_t}$$

我们有

$$E_t[S_T] = S_t e^{r(T-t)} \tag{11.53}$$

⊖ 关于 t 对(11.48)两边求导数．

所以

$$S_0 = \mathrm{e}^{-rT} E_0[S_T] \tag{11.54}$$

也就是说，当 $t=0$ 时，资产价格等于预期的未来价格以折现率 r 折现．对任意时间 t，我们相应地就有

$$S_t = \mathrm{e}^{-r(T-t)} E_t[S_T] \tag{11.55}$$

我们再回顾一下伊藤引理在这里的应用．通过伊藤引理，我们能够得到关于 Z_t 线性的积分方程(11.47)．这样我们就能将一个非线性函数的期望求值问题转化为一个线性函数的期望求值问题．再通过维纳过程增量均值为 0 的性质，我们可以将含随机变量的积分消去．而第二个积分只是关于时间的，因而可以用常规微积分中的方法来处理． 187

如果我们在这里使用的不是伊藤引理，而是常规的微积分，方程(11.43)将成为

$$\mathrm{d}Z_t = \sigma \mathrm{e}^{\sigma W_t} \mathrm{d}W_t \tag{11.56}$$

这时，我们所得到的股价期望值是

$$E[S_T] = S_t \mathrm{e}^{\left(r+\frac{1}{2}\sigma^2\right)(T-t)} \tag{11.57}$$

因此，我们可以看出来，如果使用常规微积分，当前股票价格不等于预期的未来价格以 r 为折现率的折现——鞅等式在这里就不成立了．

11.4 随机微分方程的主要模型

有一些随机微分方程模型在现实的金融市场中也相当有用．本节我们将讨论这些模型，并介绍它们可以用来描述哪些资产价格，以及它们因何而有用．

11.4.1 线性常系数随机微分方程

我们首先介绍一类最简单的随机微分方程．在这类方程中，漂移和扩散系数与随时间变化的信息无关：

$$\mathrm{d}S_t = \mu \mathrm{d}t + \sigma \mathrm{d}W_t, \quad t \in [0, \infty) \tag{11.58}$$

其中 W_t 是方差为 t 的标准维纳过程．

在这样的随机微分方程中，μ 和 σ 没有下标 t，这表示它们都是常数．所以，它们不依赖于信息集 I_t. ΔS_t 在长度为 h 的小区间上的均值为

$$E_t(\Delta S_t) = \mu h \tag{11.59}$$

方差为

$$V[\Delta S_t] = \sigma^2 h \tag{11.60}$$

图 11-2 是该随机微分方程一个路径的实例．我们用计算机模拟得到该轨迹．首先，选取适当的 μ 和 σ：

$$\mu = 0.01 \tag{11.61}$$

$$\sigma = 0.03 \tag{11.62}$$

然后设定一个很小，但是有限的时间间隔

$$h = 0.001 \tag{11.63}$$

这是对无穷小间隔 dt 的一种近似．起始点为

$$S_0 = 100 \tag{11.64}$$

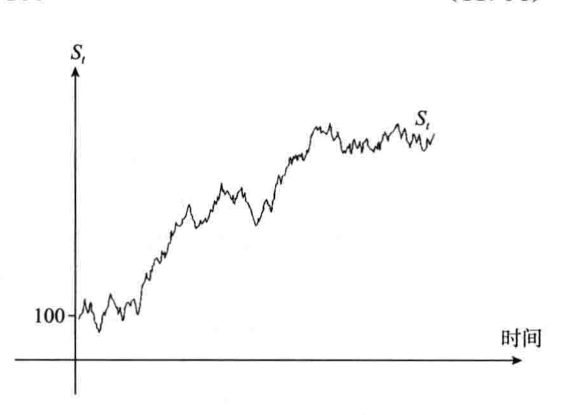


图 11-2 线性常系数随机微分方程的一条样本路径

最后，用随机数生成器产生 1000 个独立的正态分布变量．它们均值为 0，方差为 0.001．注意在这里，是因为(11.58)中的 W_t 是鞅，我们才可以使用这些独立的(正态分布)随机变量来模拟这个价格过程在小区间上的增量．

188

通过对资产价格增量的离散化近似，以及下面这个迭代公式，我们可以得到图 11-2 中的 S_t：

$$S_k = S_{k-1} + 0.01(0.001) + 0.03(\Delta W_k), \quad k = 1,2,\cdots,1000 \tag{11.65}$$

其中 S_0 已知，只需不断用正态随机数作为 ΔW_t，即可逐个得到 S_k．

从图中可以看到，S_t 看起来像在斜率为 μ 的直线附近波动．σ 的大小决定了这种波动偏离直线的程度．值得注意的是，这个价格过程的波动幅度并不会随着时间推移而增加．

这个随机微分方程模型与我们的某些现实经验是相符的．具体地说，如果一项资产的价格过程关于时间是稳定的，即趋势是线性的，而方差不会增加，那选择这样的随机微分方程来建模是非常合适的．最后，当使用这个随机微分方程建模时，我们要确定资产价格过程不存在系统性的跳跃现象．

11.4.2 几何随机微分方程

一般我们用来对资产价格建模的随机微分方程并不是上节中介绍的线性常系数随机微分方程过程，而是用几何过程．这类模型最早是由 Black 和 Scholes 开始使用与研究的：

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \quad t \in [0,\infty) \tag{11.66}$$

对于该模型

$$a(S_t,t) = \mu S_t \tag{11.67}$$

且

$$\sigma(S_t,t) = \sigma S_t \tag{11.68}$$

由此可以看出，漂移和扩散系数依赖于 t 时刻的信息．这种依赖关系非常简单与直接：漂移率和标准差与 S_t 成比例变化．如果我们将(11.66)两边同时除以 S_t，可以得到

$$\frac{dS_t}{S_t} = \mu S_t + \sigma dW_t \tag{11.69}$$

这说明尽管价格变化量的漂移项和扩散项会随着时间变化，但是价格变化率的漂移项和扩散项都是不随时间变化的常数．

图 11-3 是我们通过对

$$dS_t = 0.15 S_t dt + 0.30 S_t dW_t \tag{11.70}$$

做有限差分近似后所得到的 S_t，起点 $S_0 = 100$. 从图中可以看到，S_t 由两部分组成．首先，这个价格过程存在增速为15%的指数增长趋势．其次，围绕该趋势存在随机波动，而且随着时间推移，由于价格变得更高，方差也会越来越大．那么，在实际生活中，这个模型与常系数随机微分方程之间该如何选择呢？

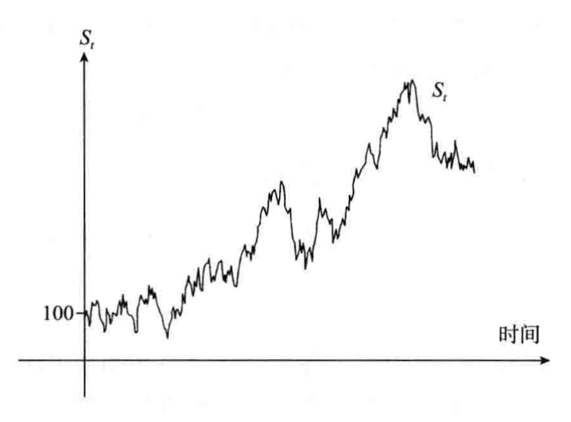


图 11-3 几何随机微分方程的一条样本路径

我们已经看到常系数随机微分方程围绕线性趋势波动，而几何过程模型围绕指数型的增长趋势波动．而对于大多数资产价格而言，指数型增长，也即是几何过程所描述的增长趋势，更加贴合实际．

那么，我们在几何过程模型中所做的关于扩散系数的假设，即它与价格的比例为常数，是否也更贴近实际呢？

189

我们首先注意到，S_t 在 t_k 和 t_{k+1} 之间增量的"方差"可以近似为

$$V[S_k - S_{k-1}] = \sigma^2 S_{k-1}^2 \tag{11.71}$$

因此，在几何过程的模型中，价格增量的方差与价格本身的平方成正比．在某些实际情况下，这种方法高估了 S_t 的方差．

11.4.3 平方根过程

另一个与刚才所介绍的模型相似的是平方根过程模型

$$dS_t = \mu S_t dt + \sigma \sqrt{S_t} dW_t, \quad t \in [0, \infty) \tag{11.72}$$

在这个模型中，S_t 的增长服从指数趋势，而标准差则是 S_t 平方根的函数，不再是 S_t. 这使得误差项的方差与 S_t 的比例为常数．

因此，如果资产价格的波动不会随 S_t 的增加而大幅增加，用这样的模型更适合．当然，条件是 $S_t > 1$.

图 11-4 是我们给出的一个例子．其中 dW_t 与图 11-3 中使用的值相同．我们考虑方程

$$dS_t = 0.15 S_t dt + 0.30 \sqrt{S_t} dW_t \tag{11.73}$$

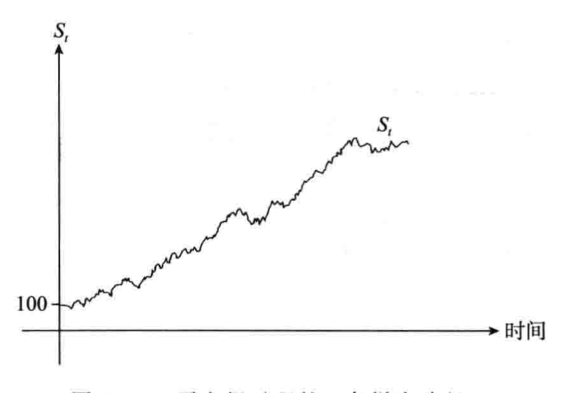


图 11-4 平方根过程的一条样本路径

其中漂移和扩散系数与图 11-3 的模型中相同，而扩散项不再与 S_t 成正比．我们选择的

初始值也同样是 $S_0=100$.

显然，图 11-4 中的波动明显比图 11-3 更缓和，而两者的总体趋势是相同的.

最后，我们要提醒读者，在这个模型中，参数 σ 不能再被看作 S_t 的波动率. 而按照金融市场的惯例，人们往往是根据资产价格的波动率来进行报价的.

11.4.4 均值回归过程

均值回归模型是一个在实践中非常有用的资产价格模型：[⊖]

$$dS_t=\lambda(\mu-S_t)dt+\sigma S_t dW_t \tag{11.74}$$

当 S_t 小于某个"均值"μ，括号中的项$(\mu-S_t)$将为正. 这使得 dS_t 也更加可能取正值. 在这个模型中，S_t 最终将趋于 μ 并在 μ 附近波动.

还有一种与之类似的随机微分方程，其漂移项也是均值回归型，而方差项依赖于 S_t 的平方根：

$$dS_t=\lambda(\mu-S_t)dt+\sigma\sqrt{S_t}\,dW_t \tag{11.75}$$

均值回归的随机微分方程模型与之前我们介绍的模型有一个重要区别.

均值回归过程也有一个总体趋势，但它对趋势的偏离并不是完全随机的. 尽管过程 S_t 可能会暂时脱离模型所描述的趋势，但它最终还是会回归这个趋势. 当然，这种脱离有可能会持续比较长的时间，而这段脱离趋势的时间的平均长度则是由参数 λ 控制($\lambda>0$). λ 越小，脱离趋势的时间就可能越长. 因此，资产价格可能表现出某种可预测的周期性，而这常常使得模型与市场有效性产生矛盾.

190

图 11-5 给出了一个均值回归模型的例子. 我们取

$$\mu=0.05,\quad \lambda=0.5,\quad \sigma=0.8 \tag{11.76}$$

也就是说从长期来看，单位时间内均值为 5%、波动率为 80%，而 $\lambda=50\%$则代表在单位时间里对价格(往均值方向)调整的比例.

现在我们取时间间隔为 $h=0.001$，即在单位时间内，我们将观察到 1000 个 S_t 的值.

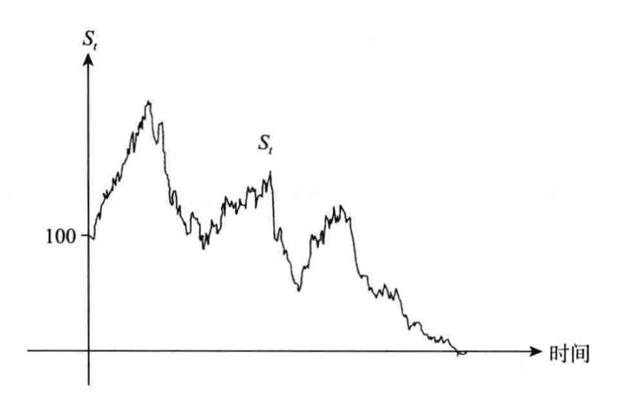


图 11-5　均值回归过程的一条样本路径

在得到均值为 0、方差为 0.001 的正态分布随机数后，我们可以通过下面的增量迭代，得到 S_t 的样本路径

$$\Delta S_k=0.5(0.05-S_{k-1})(0.001)+0.8\Delta W_k,\quad k=1,2,\cdots,1000 \tag{11.77}$$

⊖ 该模型常用于利率模型.

其中初始点位 $S_0=100$.

图 11-5 给出了这样的一个样本路径．由于本例中的扩散项不依赖于 S_t，价格可能出现负值.

11.4.5 Ornstein-Uhlenbeck 过程

再介绍一个有用的随机微分方程模型，其价格对应的过程是 Ornstein-Uhlenbeck 过程

$$dS_t = -\mu S_t dt + \sigma dW_t \tag{11.78}$$

其中 $\mu>0$. 这里，漂移项通过参数 μ 与 S_t 负相关，而扩散项具有常数参数．显然，这是“均值回归随机微分方程”的一个特例．

该模型可以用来描述价格在 0 附近波动的情形．这个价格过程可能在一段时间里脱离趋势，但最终还是会回归长期均值 0. 而参数 μ 控制这种脱离的平均时长：μ 越大，S_t 越快回归零点．

11.5 随机波动率

在上面介绍的所有随机微分方程中，我们都是用常规方法对漂移和扩散系数进行建模. 最简单的是常系数模型，最复杂的是均值回归模型．

对更加一般化的随机微分方程而言，漂移和扩散参数都是随机的．而在金融衍生品定价中，这是很有意义的，因为它意味着已知 S_t 时，波动率不仅随时间变化，而且是随机的.

例如，考虑资产价格 S_t 的随机微分方程

$$dS_t = \mu dt + \sigma_t dW_{1t} \tag{11.79}$$

其中漂移参数是常数，而扩散参数是随时间变化的．更具体地说，σ_t 服从一个偏微分方程

$$d\sigma_t = \lambda(\sigma_0 - \sigma_t)dt + \alpha\sigma_t dW_{2t}$$

其中维纳过程 dW_{1t}、dW_{2t}很可能是不独立的．

该方程描述了关于波动率的信息：在这里，资产波动率具有长期均值 σ_0. 但在任意时间 t，真实波动率可能偏离长期均值，调整参数为 λ . dW_{2t}是波动率的不可预测波动，与资产价格中的不可预测的波动不同 . $\alpha>0$ 为参数．

191

市场参与者必须预测资产价格和波动率．使用这种随机微分方程，我们可以得到更复杂的模型，以便更好地描述现实中的金融现象．另一方面，随机波动率增加了额外的扩散项，可能也增加了需要对冲的风险．这一定程度上使得模型不“完善”.

纯跳跃框架

在第 8 章，我们讨论了列维过程．在本章的剩余部分，我们将引入纯跳跃过程，我们将讨论如何得到纯跳跃过程以及介绍这类随机过程的基本性质．然后我们将导出价值函数的偏积分微分方程(PIDE). 之前我们讨论的模型都侧重于调整所研究过程的波动率以适应动态的波动率结构，或者调整漂移项以模拟市场中的均值回归特征．但是现实的金融市场

中，价格和各种比率并不总是连续变化的，而是会发生瞬时的跳跃．这些跳跃产生的影响在期权市场中是很常见的．实际上，专家们早已发现了在股票价格模型中引入跳跃成分的重要性．他们认为纯扩散模型难以解释短期期权定价中的非常明显的“微笑”效应．因此他们做了大量努力来将价格跳跃引入模型中，于是有了泊松型跳跃项和跳跃-扩散模型等模型.

过去的金融数学中的模型往往都以布朗运动为基础，且具有连续的价格路径，而现在跳跃过程在金融各个领域都起着非常重要的作用(见 Carr and Madan，1999)．需要引入这一类新过程的原因是原来的扩散项分布性质很多时候都与实际经验不符合．另一个原因是由于半鞅理论和计算机能力的发展，人们对跳跃过程的理解和使用都有了巨大进步．

最简单的跳跃过程是单跳跃过程．设 T 为关于某信息结构的停时，该信息结构由信息流$(F_t)_{t\geqslant 0}$表示，那么

$$X_t = 1_{T\leqslant t} \quad (t \geqslant 0) \tag{11.80}$$

值为 0，直到某事件发生，然后值为 1. 该过程看起来很简单，但是我们可以用这个过程来描述一个公司信用违约时间，因而它在信用风险建模中非常重要．我们将在第 23 章中进一步介绍该模型．

现考虑几何布朗运动，随机微分方程为：

$$\frac{\mathrm{d}S_t}{S_t} = \mu \mathrm{d}t + \sigma \mathrm{d}W_t \tag{11.81}$$

其中 μ 是单位时间的瞬时期望回报，σ 是单位时间的瞬时波动率，W_t 是标准维纳过程．主要有两种方法能在这个框架中加入跳跃．第一种是直接加入一个跳跃项，构成所谓的跳跃一扩散过程(Merton，1976)．第二种是对这个过程做时间变换，使维纳过程沿不同概念下的时间推移，而不是标准的日历时间，从而我们就可能使过程产生跳跃．对于这样的过程，我们称之为纯跳跃过程(Madan 等，1998；Carr 等，2002)．

192

为了扩展到跳跃一扩散的基本框架，我们做如下介绍．假设 N_t 是一个独立于跳跃规模 J 和维纳过程 W_t 的泊松过程，在测度 P 下每单位时间的发生概率为λ，则其增量满足

$$\mathrm{d}N_t = \begin{cases} 1 & \text{概率为 } \lambda \mathrm{d}t \\ 0 & \text{概率为 } 1-\lambda \mathrm{d}t \end{cases}$$

而跳跃规模的期望为

$$\kappa = E(\mathrm{e}^J - 1) \tag{11.82}$$

每次跳跃的大小均来自独立同分布的随机变量，且我们事先就已经设定好了其概率分布．那么，价格的动态变化过程可表示为方程

$$\frac{\mathrm{d}S_t}{S_t} = (\mu - \lambda\kappa)\mathrm{d}t + \sigma \mathrm{d}W_t + (\mathrm{e}^J - 1)\mathrm{d}N_t \tag{11.83}$$

该模型叫做 Merton 跳跃-扩散模型(Merton，1976)．在风险中性测度下我们可以将这个方程写为

$$\frac{\mathrm{d}S_t}{S_t} = (\mu - \lambda^* \kappa^*)\mathrm{d}t + \sigma \mathrm{d}\widetilde{W}_t + (\mathrm{e}^J - 1)\mathrm{d}N_t \tag{11.84}$$

方差伽马过程

方差伽马(VG)过程是纯跳跃过程，常用来描述高活跃性、服从正态分布、在任一时间段内有无穷多次跳跃的事件．和其他许多跳跃模型不同，VG 过程模型不一定含有扩散项，Black-Scholes 模型是一个参数化且用于描述高活跃性事件的特例．和一般的扩散过程不同，如果对 VG 模型中的价格取对数，并对这些对数价格的变化量取绝对值，那么所有这样的绝对值的和值是有限的．由于 VG 具有有限变差，它可以写为两个递增过程的差，前一个递增过程描述了价格的上升，后一个则描述价格的下降．而这两个递增过程本身是伽马过程．

1. 随机微分方程

VG 过程是一个三参数的广义布朗运动，常用来对一些市场变量的对数变化规律建模. VG 过程可以按以下方法得到：对一个漂移参数和波动率均为常数的布朗运动做随机时间变换，而这个时间变换则是由伽马过程所产生的．即

$$b(t,\sigma,\theta) = \theta t + \sigma W_t$$

$$X(t;\sigma,\upsilon,\theta) = b(\gamma(t;1,\upsilon),\sigma,\theta) = \theta\gamma(t;1,\upsilon) + \sigma W(\gamma(t;1,\upsilon))$$

我们可以这样理解这个模型：时间轴上的单位时间长度都被转换成一个具有经济意义的时间长度，而每一个这样的时间长度都是由服从均值为 1 单位、方差大于零的伽马分布 $\gamma(t;1,\upsilon)$ 的独立随机变量所确定．因此，如果一个市场中，不同的时间段中有不同的交易活跃性，那么我们就可以用这个模型来描述．根据 Carr 等人(2003)的讨论，这里的随机时间变换背后的经济意义接近于布朗运动做标度变换后产生的随机波动率．这个性质将幅度的变化和时间的变化联系起来，那么，波动率的随机变化也能够通过时间的随机变化捕捉到．因此 VG 模型中的随机时间变换是一种用来描述纯跳跃过程中的随机波动率的方法.

193

在 VG 模型中，我们知道随机时间的长度是一个服从伽马分布的随机变量，而单个时间段内的回报值(按连续复利计算)关于这个随机时间长度的条件分布是正态分布．最终的过程和对应的定价模型为我们提供的是一个稳定的三参数广义标准布朗运动模型．VG 模型下资产价格过程的对数由以下方程给出

$$\ln S_t = \ln S_0 + (r - q + \omega)t + X(t;\sigma,\upsilon,\theta)$$

或者等价地，

$$S_t = S_0 \mathrm{e}^{(r-q+\omega)t + X(t;\sigma,\upsilon,\theta)}$$

ω 是待定的系数，它的取值由下面这个等式决定

$$E(S_t) = S_0 \mathrm{e}^{(r-q)t}$$

在 VG 模型下，t 时刻资产价格的对数关于伽马时间变化 g 的条件密度函数是正态密度函

数. 非条件密度函数可以试着通过积分将 g 处理掉从而得到 .

$$f(x;\sigma,\upsilon,\theta)=\int_0^{\infty}\phi(\theta g,\sigma^2 g)\times \text{gamma}=\left(\frac{t}{\upsilon},\upsilon\right)\mathrm{d}g\int_0^{\infty}\frac{1}{\sigma\sqrt{2\pi g}}\exp\left(-\frac{(x-\theta g)^2}{2\sigma^2 g}\right)\frac{g^{t/\upsilon-1}\mathrm{e}^{-g/\upsilon}}{\upsilon^{t/\upsilon}\Gamma(t/\upsilon)}\mathrm{d}g$$

在一般化的该模型中，参数不但可以控制布朗运动的波动率，还可以控制(ⅰ)尖峰厚尾性，即左右尾部概率相比正态分布较高；(ⅱ)偏度，即左右对称程度 .

VG 模型还有一个引人注目的特性，对数正态的密度函数与 Black-Scholes 公式都是它的一个参数化特例 .

2. 特征函数

现在我们来计算 VG 过程的特征函数 . 首先我们计算以伽马时间 g 为条件的条件特征函数：

$$E(\mathrm{e}^{iuX_t}\mid g)=E(\mathrm{e}^{iu(\theta g+\sigma W_g)})=\mathrm{e}^{iu\theta g}E(\mathrm{e}^{iu\sigma\sqrt{g}Z})=\mathrm{e}^{iu\theta g}\,\mathrm{e}^{-\frac{(u\sigma\sqrt{g})^2}{2}}=\mathrm{e}^{iu\theta g}\,\mathrm{e}^{\frac{-u^2\sigma^2 g}{2}}=\mathrm{e}^{i\left(u\theta+i\frac{u^2\sigma^2}{2}\right)g}$$

现计算 VG 过程的特征函数，我们要对 g 积分 .

$$E(\mathrm{e}^{iuX_t})=E_g\left(\mathrm{e}^{i\left(u\theta+i\frac{u^2\sigma^2}{2}\right)g}\right)=\int_0^{\infty}\mathrm{e}^{iu\theta g}\,\mathrm{e}^{\frac{-u^2\sigma^2 g}{2}}\frac{g^{t/\upsilon-1}\mathrm{e}^{-g/\upsilon}}{\upsilon^{t/\upsilon}\Gamma(t/\upsilon)}\mathrm{d}g$$

这就是形状参数为 $\frac{t}{\upsilon}$ 规模参数为 υ 的伽马过程的特征函数在 $\mu\theta+i\frac{\mu^2\sigma^2}{2}$ 的值 . 结合(5.43)，我们可得

$$E_g\left(\mathrm{e}^{i\left(u\theta+i\frac{u^2\sigma^2}{2}\right)g}\right)=\left(\frac{\frac{1}{\upsilon}}{\frac{1}{\upsilon}-i\left(u\theta+i\frac{u^2\sigma^2}{2}\right)}\right)^{\frac{t}{\upsilon}}=\left(\frac{1}{1-iu\theta\upsilon+\frac{u^2\sigma^2\upsilon}{2}}\right)^{\frac{t}{\upsilon}} \tag{11.85}$$

综上所述，我们得到 t 时刻参数为 σ、υ 和 θ 的 VG 过程的特征函数为

194

$$E(\mathrm{e}^{iuX(t)})=\left(\frac{1}{1-iu\theta\upsilon+\sigma^2u^2\upsilon/2}\right)^{\frac{t}{\upsilon}} \tag{11.86}$$

11.6 小结

本章引入了随机微分方程解的概念 . 我们介绍了两类解的区别 . 强解与常微分方程的解类似 . 弱解相对比较新颖 .

这里我们没有具体讨论弱解 . 我们将在之后的章节中讨论一个重要例子 . 本章还介绍了资产定价模型中常用的几种随机微分方程 .

11.7 参考阅读

本章中我们介绍了部分 Oksendal(2000)的研究，他的著作中还包括了其他一些随机微分方程 . 关心应用层面的读者可以阅读随机微分方程数值解方面的文献 . Kloeden 等(1994)非常易懂和全面 . 可以说，学习随机微分方程最好的方法是研究它们的数值解 .

11.8 习题

1. 考虑以下随机微分方程：

$$d(W_t^3) = 3[W_t dt + W_t^2 dW_t]$$

(a) 写出该随机微分方程对应的积分方程．

(b) 求以下积分的值

$$\int_0^t W_s^2 ds$$

2. 考虑几何随机微分方程：

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

其中 S_t 代表某资产指数．该指数的现值为 $S_0=940$. 已知年波动率为 0.15. 无风险利润率为常数 5%. 此外，该指数的计算不受红利影响．你所关心的是时长为 8 天的时间段，你可以将这段时间均分为 4 等份，每小段为 2 天，记为 Δ.

(a) 用抛硬币的方法产生随机误差，作为 dW_t，其中

$$H = +1, \quad T = -1$$

(b) 如果 $\Delta \to 0$，如何确保抛硬币产生的随机过程 dW_t 的均值和方差与题目所要求的相匹配？

(c) 生成 3 个总时长为 8 天的 S_t 的随机路径．

3. 考虑以下描述某证券价格的线性随机微分方程：

$$dS_t = 0.01 S_t dt + 0.05 \sigma S_t dW_t$$

已知 $S_0=0$. 假设现有以该证券为标的的欧式看涨期权，期限 $T=1$，执行价格 $K=1.5$. 假设无风险利率为 3%.

(a) 利用你的计算机，产生 5 个均值为 0、方差为$\sqrt{2}$的随机变量．

(b) 取 $\Delta t=0.2$，利用上一小题中所产生的随机变量作出 S_t 的拟合轨迹．

(c) 求看涨期权在到期日的价值．

(d) 再取 5 个同分布随机数，重复以上试验．

(e) 如果我们将该实验进行 1000 次，其中任意两次试验得到的期权价格会有明显差别吗？为什么？

(f) 我们能否将两份蒙特卡洛样本结合起来，构成 2000 条路径，来计算看涨期权价格？ 195

4. 考虑以下随机微分方程：

$$dS_t = 0.05 dt + 0.1 dW_t$$

其中 dW_t 由以下过程近似：

$$\Delta W_t = \begin{cases} +\Delta & \text{概率为 } 0.5 \\ -\Delta & \text{概率为 } 0.5 \end{cases}$$

(a) 假设 $S_0=100$，设区间间隔为 $\Delta=1$. 计算从 $t=0$ 到 $t=3$，S_t 的值．

(b) 令 $\Delta=0.5$，重复上述计算．

(c) 画出以上两个 S_t 的轨迹图．

(d) 假如 $\Delta=0.01$，这种图看起来是什么样的？

(e) 将 S_t 的方差乘以 3，令 $\Delta=1$，得到一个新的 S_t.

(你可以选择通过抛硬币的方法来得到随机数．)

5. 考虑均值回归过程 $dS_t=\lambda(\mu-S_t)dt+\sigma StdW_t$. 建立 $\Delta=\frac{1}{252}$的时间分割，假定月模型参数为：$\lambda=0.75$，$\sigma=0.35$，$S_0=40$. 考虑 $\mu=30$ 和 $\mu=50$ 两种情况．以此为基础，建立看涨期权定价的蒙特卡洛算法．股票价格可以按以下步骤计算：

$$S_{t+1}=S_t+dS_t \tag{11.87}$$

$$dS_t=\lambda(\mu-S_t)\Delta+\sigma S_t W \tag{11.88}$$

这里 W 是均值为 0、方差为 Δ 的正态分布的随机变量．

生成 1000 条 3 年期的样本路径．运用蒙特卡洛法计算执行价格为 30，40，50，60，70 时看涨期权的价格．说明当改变 μ 后期权价格将如何变化．

6. 编写欧式看涨期权定价的模拟程序，标的股票价格服从波动率 $\sigma=0.1$ 的几何布朗运动．其余参数为 $r=0.05$，$q=0$，$S=K=10$，$T=1$.

196

第 12 章　衍生品定价：偏微分方程

12.1　引言

截至目前我们已经学习了连续时间下随机过程建模的主要工具，以及在这些情况下如何(或者无法)求导数和积分．

我们并不是单纯地为了学习这些工具而介绍它们，而是因为它们在金融市场上各种衍生品定价中非常有用．与纯理论研究不同，这些方法已经成为了金融市场上专业人士的常用工具．事实上，由于衍生品的一些特殊性质，抽象理论模型在这个领域中与实际问题的契合程度远比金融研究中的其他领域要好．

现代金融已发展出了两种主要的衍生资产定价方法．第一种是利用偏微分方程，即本章的内容．第二种需要将所研究的过程转化为鞅．这要用到等价鞅测度，是第 14 章的内容．理论上，这两种方法将得出同样的答案．在应用中，根据问题的特点，可能其中某一种方法更为简便．不过这两种定价方法背后的数学工具是迥异的．

首先，我们将简要介绍为什么我们可以将证券定价转化为偏微分方程问题．这些结果将在第 13 章用到．

12.2　建立无风险投资组合

衍生工具是建立在证券上的合约，这些合约具有有限的期限．在到期日 T，其价格 F_T 仅与标的资产价格 S_T 和时间 T 有关：

$$F_T = F(S_T, T) \tag{12.1}$$

这意味着我们知道 $F(S_T, T)$的确切形式(值)．假设这种关系在除到期日 T 以外的其他时刻也成立，那么衍生品价格可以写作

$$F(S_t, t) \tag{12.2}$$

该价格的变化量记作 $\mathrm{d}F_T$．从表面上看，市场参与者不知道 $F(S_t, t)$除到期日之外的确切形式．我们需要找出该函数．

这意味着如果我们知道 S_t 过程的运动规律——例如，如果我们知道一个确定 $\mathrm{d}S_t$ 运动方式的方程——那么我们就能通过伊藤引理得到 $\mathrm{d}F_T$．但这表示增量 $\mathrm{d}F_T$ 和 $\mathrm{d}S_t$ 的风险来源是相同的，即 $\mathrm{d}S_t$ 的干扰项．也就是说，至少在目前的例子中，$\mathrm{d}F_T$ 和 $\mathrm{d}S_t$ 依赖于同一个干扰项．这种依赖关系使我们有可能建立在连续时间下的无风险的投资组合．

投资 P_t 美元于这个的一个包含 $F(S_t, t)$和 S_t 的投资组合：

$$P_t = \theta_1 F(S_t, t) + \theta_2 S_t \tag{12.3}$$

其中 θ_1、θ_2 分别是衍生工具和标的证券的量，它们是投资组合的权重．

随着时间 t 的推移，投资组合的价值会变化，因为 $F(S_t,t)$和 S_t 会变化．取 θ_1、θ_2 为常数，可以将变化值记为⊖

$$dP_t = \theta_1 dF_t + \theta_2 dS_t \tag{12.4}$$

通常，θ_1、θ_2 会随时间变化，因此也应该带有时间下标．但现在先让我们忽略这一点．在该方程中，由于 dS_t 中含有干扰项 dW_t，增量 dF_T 和 dS_t 中也都含有不可预测成分．

一个关于记号的注意点．在小区间 dt 中，dF_t 仍然要看作 $F(S_t,t)$的全微分．不要和 F_t 混起来，后者是 $F(S_t,t)$关于 t 的偏导．

198

我们主要感兴趣的是衍生品的价格及其变化过程．所以，我们首先建立描述资产价格 S_t 变化的模型，由此我们研究 $F(S_t,t)$如何变化．相应地，我们假设随机微分 dS_t 服从随机微分方程

$$dS_t = a(S_t,t)dt + \sigma(S_t,t)dW_t \tag{12.5}$$

由此，我们可以使用伊藤引理得到 dF_t：

$$dF_t = F_t dt + \frac{1}{2}F_{ss}\sigma_t^2 dt + F_s dS_t \tag{12.6}$$

用方程(12.5)将 dS_t 替换掉，得到衍生资产价格的随机微分方程：

$$dF_t = \left[F_s a_t + \frac{1}{2}F_{ss}\sigma_t^2 + F_t\right]dt + F_s\sigma_t dW_t \tag{12.7}$$

我们将漂移系数简记为 a_t，扩散系数简记 σ_t. 如果我们知道函数 $F(S_t,t)$的具体形式，我们可以计算偏导数 F_s、F_{ss}、F_t，然后就能得到这个包含衍生品价格变化信息的随机微分方程．但是，函数 $F(S_t,t)$的形式是未知的．我们能够用以下步骤来求出它．

首先我们看(12.7)中的随机微分方程，它描述了 dF_t 的变化情况，它和 S_t 一样，都受到同一个维纳过程 dW_t 影响．理论上，我们应该可以用其中一个公式将另一个公式中的随机项替换掉．而在建立无风险投资组合过程中，我们所做的正是这一点．

下面我们介绍如何完成这一步．首先注意到，投资组合的权重 θ_1,θ_2 是由市场参与者选择的．

其次，总可以选取适当的权重，使得 dP_t 与干扰项 dW_t 相独立．原因如下．已知 dF_t 和 dS_t 含有相同的不可预测项，且存在合适的 θ_1，θ_2，能够将方程(12.4)中的 dW_t 项消去. 要达到该目的，再次考虑

$$dP_t = \theta_1 dF_t + \theta_2 dS_t \tag{12.8}$$

利用(12.6)将 dF_t 替换掉：⊜

$$dP_t = \theta_1\left[F_s dS_t + \frac{1}{2}F_{ss}\sigma_t^2 dt + F_t dt\right] + \theta_2 dS_t \tag{12.9}$$

在该方程中我们可以任意设定 θ_1 和 θ_2. 假设我们暂时不考虑 F_s 依赖于 S_t，取

⊖ 严格地说，这里给出的随机微分只在组合权重与 S_t 无关时正确．否则，右端会出现多余的项．这一点与我们后面介绍的 Black-Sholes 的基本框架关系密切．

⊜ θ_1、θ_2 与 S_t 无关，这一步在数学上才能成立．

$$\theta_1 = 1 \tag{12.10}$$

以及

$$\theta_2 = -F_s \tag{12.11}$$

这些组合权重按以上取值，则(12.9)中的 $\mathrm{d}S_t$ 会被消去，方程化简为

$$\mathrm{d}P_t = F_t \mathrm{d}t + \frac{1}{2} F_{ss} \sigma_t^2 \mathrm{d}t \tag{12.12}$$

所以，在已知信息集 I_t 的情况下，该表达式中没有随机项．对于所有时间 t，$\mathrm{d}P_t$ 是完全可预测、确定性的增量．这意味着投资组合 P_t 是无风险的．[1]

由于 P_t 无风险，它在区间 $\mathrm{d}t$ 的增值必须等于无风险投资的利润，否则就会产生套利．假设无风险利率为(常数)r，无红利支付，期望资本收益必须等于

$$rP_t \mathrm{d}t \tag{12.13}$$

若支付红利，设在每单位时间支付 δ，则必须等于

$$rP_t \mathrm{d}t - \sigma \mathrm{d}t \tag{12.14}$$

199

在这种情况中，资本收益加上红利收益还是等于无风险利率．[2]

以无红利的情况为例，由方程(12.12)和(12.13)可得

$$rP_t \mathrm{d}t = F_t \mathrm{d}t + \frac{1}{2} F_{ss} \sigma_t^2 \mathrm{d}t \tag{12.15}$$

每一项都含有 $\mathrm{d}t$，可以将它去掉，得到一个偏微分方程：

$$r(F(S_t,t) - F_s S_t) = F_t + \frac{1}{2} F_{ss} \sigma_t^2 \tag{12.16}$$

我们再将(12.15)中的 P_t 替换掉，将(12.16)改写为

$$-rF + rF_s S_t + F_t + \frac{1}{2} F_{ss} \sigma_t^2 = 0, \quad 0 \leqslant S_t, \quad 0 \leqslant t \leqslant T \tag{12.17}$$

其中衍生资产价格 $F(S_t,t)$ 简记为 F．

我们还有一个没有使用的信息．衍生品的到期日为 T，而 T 时刻标的资产价格和衍生资产价格的关系一般非常明确．也就是说到期日衍生品的价格为

$$F(S_T,T) = G(S_T,T) \tag{12.18}$$

其中 $G(\cdot)$ 是一个关于 S_T 和 T 的已知函数．例如，对于欧式看涨期权，执行价格为 K，则其到期日的价格为

$$G(S_T,T) = \max[S_T - K, 0] \tag{12.19}$$

根据该方程，在到期日，如果股票价格低于执行价格，$S_T - K$ 为负，看涨期权不会被执行，它将没有价值．反之，看涨期权的价值将等于股票价格与执行价格之差．

---

[1] 注意重要的一点：若令 θ_2 等于 $-F_s$，则 θ_2 会随时间变化．对于期权等非线性产品，F_s 将会是 S_t 的函数．这意味着这样做不满足无风险资产组合的数学条件，然而它却能够得出正确的 PDE.

[2] 注意 $\mathrm{d}t$ 的作用．必须经过一些无穷小时间才能赚取利息或得到红利．如果时间静止，不管利率 r 是多少，利息收益也始终为 0. 红利同理．

方程(12.17)是一个偏微分方程．方程(12.18)是这个偏微分方程的边界条件．

我们之所以可以用上述方法将干扰项从方程(12.4)中去掉是因为 $F(\cdot)$作为衍生工具的价格，它与其标的物的价格 S_t 有相同的不可预测成分 $\mathrm{d}W_t$. 因此，我们可以通过将 $F(\cdot)$和 S_t 按照一定权重组合起来，消掉它们共有的不可预测运动．最后，这样得到的投资组合 P_t 成为了无风险投资，我们可以将它的未来轨迹确定下来．

这种构建无风险投资组合的方法非常有启发性．从数学的角度来看，它并不令人满意. 根据正规的方法，我们需要根据市场完备性考虑一系列交易策略、资产的中性等价物等，建立自融资投资组合．Jarrow(1996)是这方面内容极好的参考读物．下一节我们将具体介绍这方面内容．

12.3 偏微分方程方法的精确性

在上一节中，给定一项资产的价格过程，我们通过一个无风险投资组合推导出了关于其衍生资产无套利价格过程的偏微分方程．

我们的基本想法是建立标的资产和它的看涨期权构成的投资组合：

$$P_t = \theta_1 F(S_t, t) + \theta_2 S_t \tag{12.20}$$

其中 θ_1、θ_2 是权重．然后我们关于 $\mathrm{d}t$ 求微分，得到

$$\mathrm{d}P_t = \theta_1 \mathrm{d}F(S_t, t) + \theta_2 \mathrm{d}S_t \tag{12.21}$$

从数学上来说，这个方程把 θ_1 和 θ_2 看作常数，因为并没有对它们求微分．到现在为止，这样做对无风险投资组合方法并没有什么不对．但是，当我们选择权重后，会发生什么呢?

200

我们取权重为：

$$\theta_1 = 1, \quad \theta_2 = -F_s \tag{12.22}$$

如同我们在上一节中介绍的那样，这样的全中可以将投资组合中不可预测的随机成分消去．但另一方面，它也违背了我们关于 θ_1、θ_2 是常数的假设．现在，θ_2 依赖于 S_t，因为通常 F_s 是 S_t 和 t 的函数．所以如果先用(12.22)中 θ_1、θ_2 的取值来替换(12.20)中的 θ_1、θ_2，然后再两边做微分，这样我们会得到和(12.21)截然不同的结果．下面我们具体来看．

将 θ_1、θ_2 的取值带入(12.22)后，得

$$P_t = F(S_t, t) - F_s(S_t, t) S_t \tag{12.23}$$

关于 t 求微分并约掉 t，得

$$\mathrm{d}P_t = \mathrm{d}F(S_t, t) - F_s \mathrm{d}S_t - S_t \mathrm{d}F_s - \mathrm{d}F_s(S_t, t)\mathrm{d}S_t \tag{12.24}$$

由于 F_s 依赖于 S_t，我们有了新出现的第三项，这一项与时间有关而且是随机的．通常情况下，这一项不为 0. 事实上，我们能使用伊藤引理并计算 $\mathrm{d}F_s$，后者是 S_t 和 t 的函数. 这相当于对衍生品的 DELTA 做随机微分：

$$\mathrm{d}F_s(S_t, t) = F_{st}\mathrm{d}t + F_{ss}\mathrm{d}S_t + \frac{1}{2}F_{sss}\sigma^2 S_t^2 \mathrm{d}t$$

这里出现了 F 的三阶导，因为在我们使用伊藤引理时，F 已经关于 S_t 求过微分了．将 $\mathrm{d}S_t$

带入，可得

$$dF_s(S_t,t)=F_{st}\,dt+F_{ss}(\mu S_t\,dt+\sigma\,dW_t)+\frac{1}{2}F_{sss}\sigma_t^2S_t^2\,dt$$

$$=[F_{st}+F_{ss}\mu S_t+\frac{1}{2}F_{sss}\sigma_t^2S_t^2]dt+F_{ss}\sigma S_t\,dW_t$$

综上所述，当 θ_2 等于 $-F_s$ 时，

$$P_t+\theta_1F(S_t,t)+\theta_tS_t \tag{12.25}$$

的微分为

$$dP_t=dF(S_t,t)-F_s\,dS_t-S_t\left[\left[F_{st}+F_{ss}\mu S_t+\frac{1}{2}F_{sss}\sigma_t^2S_t^2\right]dt+F_{ss}\sigma S_t\,dW_t\right]-F_{ss}\sigma_t^2S_t^2\,dt \tag{12.26}$$

这个资产组合通常不是自融资的，因为我们没有

$$dP_t=dF(S_t,t)-F_s\,dS_t \tag{12.27}$$

与(12.27)相比，(12.26)的右端有多余项，这些多余项不等于0，除非

$$S_t^2F_{ss}(\sigma\,dW_t+(\mu-r)dt)=0$$

而这通常不会成立．为了理解这一点，再次对(12.17)中的 Black-Scholes 偏微分方程关于 S_t 求微分，可以得到

$$F_{st}+F_{ss}rS_t+\frac{1}{2}F_{sss}\sigma^2S_t^2+\sigma^2F_{ss}S_t=0$$

用该方程可以消去(12.26)中大多数我们不想要的项．但仍然会留下：

$$dP_t=dF(S_t,t)-F_s\,dS_t-S_t[F_{ss}(\mu-r)S_t\,dt]-F_{ss}\sigma S_t^2\,dW_t \tag{12.28}$$

所以，为了保证投资组织 P_t 是自融资的，我们需要

$$S_t^2F_{ss}(\sigma\,dW_t+(\mu-r)dt)=0$$

该式通常不成立．

对于(12.28)的解释

虽然严格来说，这个无风险投资组合并不足够让人满意，它需要额外的投资，或者得到额外的收入．但这个方法能够帮我们得到正确的偏微分方程．我们该如何解读这一结果呢？ 201

这个答案就在(12.28)中我们认为是多余的那项中．虽然在真实概率测度 P 下这项的期望值非零，但如果我们选择无风险测度 Q，并定义一个新的维纳过程 W_t^*，即

$$dW_t^*=\sigma\,dW_t+(u-r)dt$$

我们将得到

$$E^P[S_t^2F_{ss}(\sigma\,dW_t+(\mu-r)\Delta)]\approx 0$$

所以在小区间内，投资组合 P_t 的额外支出(收入)期望值将为0. 可以理解为平均而言，该投资组合是自融资的．但是，这种“平均”是相对于风险中性测度 Q 而言的，而不是真实概率测度P．Musiela 和 Rutkowski(1997)在这个问题上提供更详细的阐述．

12.4 偏微分方程

我们将 $F(S_t,t)$简记为 F，用一般形式重写偏微分方程(12.17)，可得

$$a_0F + a_1F_sS_t + a_2F_t + a_3F_{ss} = 0,\quad 0 \leqslant S_t,\quad 0 \leqslant t \leqslant T \tag{12.29}$$

边界条件为

$$F(S_T,T) = G(S_T,T) \tag{12.30}$$

$G(\cdot)$是已知的.⊖

通过建立无风险投资组合来对衍生品进行无套利定价的方法，最后总是可以推导出一个偏微分方程．由于衍生证券依赖于标的资产，建立这种无风险投资组合的方法通常很直接．由于衍生品的不同，边界条件和对应的偏微分方程可能会更复杂．但是，总的来说，这样的方法都是围绕求解偏微分方程进行的．这是我们需要进一步讨论的．

我们将分几个步骤来介绍偏微分方程．

12.4.1 为什么偏微分方程是“方程”

在什么情况下(12.29)中的偏微分方程是一个方程？它的解是什么样的？

在代数中，求解方程通常是求出一个数字或者向量 x，而方程(12.29)中的未知量是一个函数．我们不知道 $F(S_t,t)$是哪种类型的函数，只知道我们对它求一些偏导，再乘以 a_i 并加起来，结果会等于 0. 而且，在时刻 $t=T$，该函数等于一个已知函数 $G(S_T,T)$——也就是说，方程的解必须满足这个边界条件．

所以，求解该偏微分方程意味着我们要找到一个函数，同时满足(12.29)和(12.30).

12.4.2 什么是边界条件

将一个函数的多个偏导数组合在一起，使该组合等于 0，我们就得到了一个偏微分方程. 边界条件是这种方程的一部分．在物理中，边界条件是某些物理现象的初始或最终状态，而偏微分方程描述了该现象的过程．

202

在金融中，边界条件的作用类似．它代表了衍生品的某些合约条款．对于不同的衍生品，边界条件也可能不同．最常见的边界条件是衍生合约的初始价值和最终价值．根据金融理论，衍生合约的价格在到期日一般会满足一定的条件．例如，期货价格和现货价格在交割日不会相差很多．对期权而言，期权价格必须满足某个方程，如(12.19). 对于折价债券，在到期日其资产价格等于 100.

如果没有边界条件，一般不可能找到函数 $F(S_t,t)$，使之满足已知的偏微分方程．此外，只要衍生品关于标的资产在到期日的关系是确定的，那么就必然会给市场参与者提供一个边界条件．

⊖ 本节 PDE 的书面记号与之前介绍的有所不同．例如，(12.29)中的 PDE 本来写作

$$a_0F(X,\ t)+a_1F_s(X_t,\ t)+a_2F_t(X,\ t)+a_3F_{ss}(X,\ t)=0,\quad X\geqslant 0,\quad 0\leqslant t\leqslant T \tag{12.31}$$

边界条件为 $F(X,t)=G(X,t)$. 与常见的方法相同，本节我们用 S_t 替代变量 X.

12.5 偏微分方程的分类

我们可以从几个不同的方面对偏微分方程分类．首先，偏微分方程可以分为线性或非线性．这取决于偏导数前面的系数．如果方程是 F 和它的偏导数的线性组合，那么它是线性偏微分方程．[⊖]

第二种分类方法是根据求导的阶数．如果方程中所有偏导数都是一阶的，那么该偏微分方程也是一阶的．如果存在混合偏导，或者二阶偏导，那么该偏微分方程是二阶的．对于非线性金融衍生品如期权或者包含期权的工具，它们的偏微分方程都是二阶的．

以上分类方法都与常微分方程类似．第三种分类方法是偏微分方程特有的．偏微分方程可以被分为椭圆型、抛物型、双曲型．我们在金融中遇到的偏微分方程通常都类似于抛物型偏微分方程．

我们首先研究线性一阶、二阶偏微分方程．这些例子与金融没有直接联系．但是他们可以帮助我们认识到偏微分方程是什么、边界条件为什么很重要．

12.5.1 例 1：一阶线性偏微分方程

考虑关于 $F(S_t,t)$ 的偏微分方程：

$$F_t + F_s = 0,\quad 0 \leqslant S_t,\quad 0 \leqslant t \leqslant T \tag{12.32}$$

根据该偏微分方程，$F(\cdot)$ 关于 t 偏导的相反数等于其关于 S_t 的偏导．假如 t 代表时间、S_t 代表标的证券，那么(12.32)意味着：固定 S_t，时间变化一小段导致价格发生的变化等于固定 t，S_t 变化一小段导致价格所发生变化的相反数．

在金融市场中，没有理由说这两个偏导数存在以上关系．但假设(12.32)确定下来了并且需要求解，$F(S_t,t)$ 会是一个什么样的函数呢？

我们可以直接猜测一个解：

$$F(S_t,t) = \alpha S_t - \alpha t + \beta \tag{12.33}$$

其中 α、β 是常数．根据该函数，偏微分为：

$$\frac{\partial F}{\partial t} = -\alpha \tag{12.34}$$

和

$$\frac{\partial F}{\partial S} = \alpha \tag{12.35}$$

203

显然，它们的和等于零．这说明 F 满足(12.32).

解函数(12.33)可以在三维空间中表示一个平面．如果没有边界条件，我们仅能得出这样的结论．我们不能确定 $F(S_t,t)$ 到底具体代表哪一个平面．因为我们不能确定 α 和 β 的值．我们只知道：在 $t=0$，$S_0=0$ 时，截距为 β；固定 S_t，$F(S_t,t)$ 的图像为斜率为 $-\alpha$ 的

⊖ 这意味着偏导数的系数不是 F 的函数．

直线；固定 t，图像为斜率为 α 的直线．

图 12-1 和图 12-2 展示了两个(12.32)中偏微分方程的"解". 图 12-1 是以下平面的图像：

$$F(S_t,t)=3S_t-3t+4,\quad -10\leqslant t\leqslant 10,\quad -10\leqslant S_t\leqslant 10 \tag{12.36}$$

204

此时 $F_s=3$，$F_t=-3$. 因此该函数满足(12.32)中的偏微分方程．其函数值关于 S_t 递增，关于 t 递减．

图 12-2 则是另一个例子，该函数为

$$F(S_t,t)=-2S_t+2t-4,\quad -10\leqslant t\leqslant 10,\quad -10\leqslant S_t\leqslant 10 \tag{12.37}$$

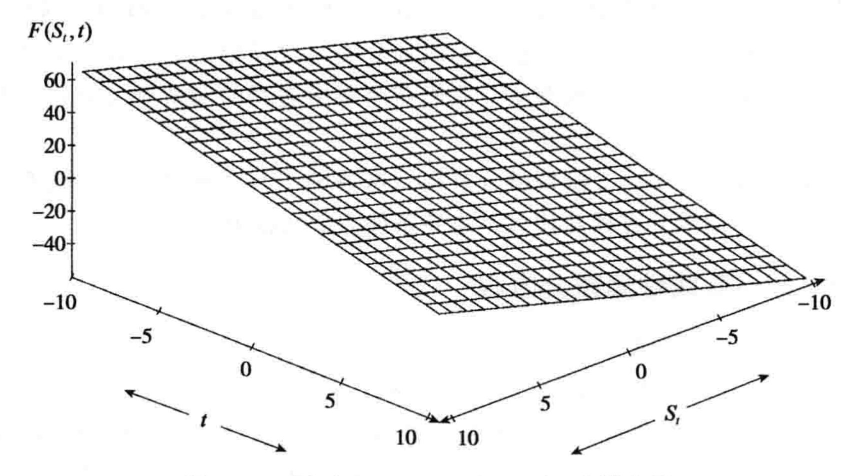


图 12-1 平面 $F(S_t,t)=3S_t-3t+4$ 的散点图

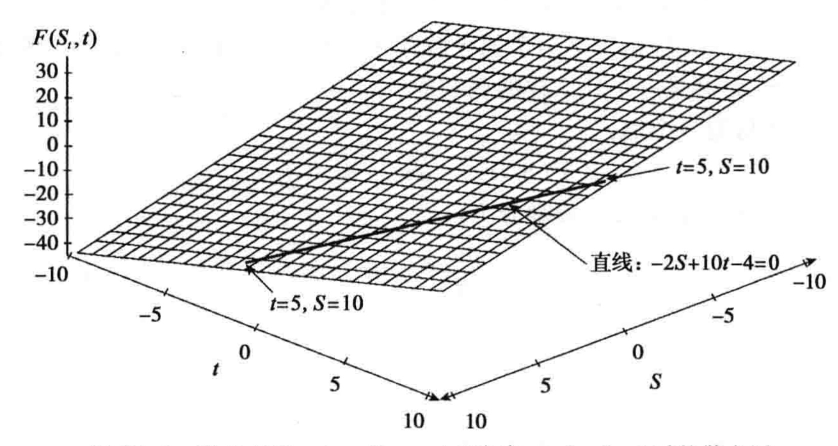


图 12-2 平面 $F(S_t,t)=\alpha S_t-\alpha t+\beta$ 中当 $\alpha=2$，$\beta=4$ 时的散点图

$F(S_t,t)$也是一个平面．但此时函数值关于 t 递增，关于 S_t 递减．它的轮廓也是直线．

式(12.36)和式(12.37)中的两个 $F(S_t,t)$完全不同，但它们都是(12.32)中偏微分方程的解. 这是因为方程(12.32)中的信息并不足以让我们能够确定 $F(S_t,t)$的具体形式．事实上，存在无数多个函数满足(12.32).

现在，如果除了(12.33)我们还知道某些边界条件，那么我们就能确定 $F(S_t,t)$ 的具体形式．例如，假设我们知道当到期日 $t=5$ 时(t 的边界)，有

$$F(S_5,5)=6-3S_5 \tag{12.38}$$

我们就能确定(12.33)中未知的 α 和 β：

$$\alpha=2 \tag{12.39}$$

$$\beta=4 \tag{12.40}$$

即图 12-2 中的平面．

另一方面，如果我们还有一个边界条件，当 $S_t=100$ 时，

$$F(100,t)=5+0.3t \tag{12.41}$$

那么就不存在有意义的解，因为方程(12.41)和(12.42)会分别得到不同的 α 和 β.

所以，如果 $F(S_t,t)$ 是平面，我们只需要一个边界条件，就能确定偏微分方程的解．

从几何上很容易理解这一点，因为一个边界条件意味着先选择一个 t(或 S_t)的端点，然后得到该平面和一个与时间轴垂直的平面的交线．在图 12-2，$t=5$ 处的边界条件

$$F(S_5,5)=6-3S_5 \tag{12.42}$$

已标记出来．注意到，图 12-1 中的候选解当 $t=5$ 时不会经过该直线，因此它不是一个解．

同理，只要 $F(S_t,t)$ 是平面，其关于 t 或 S_t 终端条件也是直线．

注意事项

对于偏微分方程

$$F_t+F_s=0 \tag{12.43}$$

其解的类型不只局限于线性解．实际上，考虑函数

$$F(S_t,t)=e^{\alpha S_t-\alpha t} \tag{12.44}$$

该函数也满足(12.43). 有边界条件才能得到唯一解．

12.5.2 例 2：二阶线性偏微分方程

我们很容易猜出例 1 中一阶线性偏微分方程的解的形式．现在我们考虑一个二阶偏微分方程：

$$\frac{\partial^2 F}{\partial t^2}=0.3\,\frac{\partial^2 F}{\partial S_t^2} \tag{12.45}$$

或者，简写为

$$-0.3F_{ss}+F_{tt}=0 \tag{12.46}$$

首先，注意到这还是一个线性偏微分方程，它是偏导数乘以常系数后的组合．

现不管边界条件，我们看看能否猜出(12.46)的解．可以看出，$F(S_t,t)$ 关于 t 的二阶导是它关于 S_t 的二阶导的 0.3 倍．这个关系在所有 S_t 和 t 处都成立．那这个偏微分方程的解函数会是什么样的呢？

205

考虑公式

$$F(S_t,t)=\frac{1}{2}\alpha(S_t-S_0)^2+\frac{0.3}{2}\alpha(t-t_0)^2+\beta(t-t_0) \tag{12.47}$$

其中 S_0、t_0 是未知常数，而参数 α 和 β 也是未知的．

现在，如果我们对 $F(S_t,t)$ 求二阶导：

$$\frac{\partial^2 F}{\partial t^2}=0.3\alpha \tag{12.48}$$

$$\frac{\partial^2 F}{\partial S^2}=1\alpha \tag{12.49}$$

以上二阶偏导数显然满足方程(12.45)．所以(12.47)中的 $F(S_t,t)$是偏微分方程(12.45)的一个解．

注意到对于固定的 $F(S_t,t)$，

$$F(S_t,t)=F_0 \tag{12.50}$$

的周线是椭圆．[⊖]

同样，方程(12.45)的解不是唯一的．因为对于任意 α、β、S_0、t_0，只要 $F(S_t,t)$满足(12.47)中的形式，它就是(12.45)的解．

若给定边界条件为，在 $S_t=10$ 处：

$$F(10,t)=100+t^2 \tag{12.51}$$

该函数在(F,t)平面上的轨迹是抛物线．

然而这样一个边界条件不足以确定所有参数 α、β、S_0、t_0．我们还需要另一个边界条件．例如，在 $t=0$ 处：

$$F(S_0,0)=50+S_0^2 \tag{12.52}$$

该方程也是一个抛物线，但这是在平面(F,S_t)上的抛物线．

图 12-3 给出了这样一个 $F(S_t,t)$的图像．它是以下函数的三维图像：

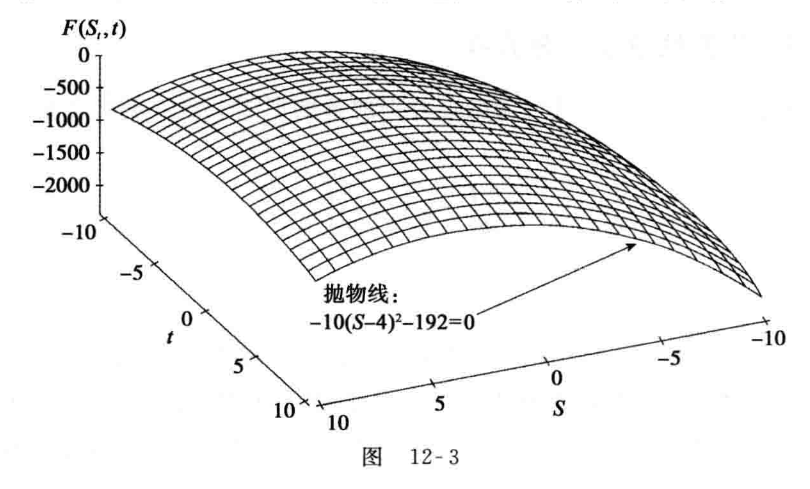


图 12-3

⊖ 参见下一节．

$$F(S_t, t) = -10(S_t - 4)^2 - 3(t-2)^2, \quad -10 \leqslant t \leqslant 10, \quad -10 \leqslant S_t \leqslant 10 \tag{12.53}$$

这个曲面的周线是椭圆．对于边界条件，我们取 $t=10$ 作为 t 的终点值，取抛物型边界条件： 206

$$F(S_{10}, 10) = -10(S_{10} - 4)^2 - 192 \tag{12.54}$$

$S_t=0$ 处的边界条件也是一个抛物线：

$$F(0, t) = -160 - 3(t-2)^2 \tag{12.55}$$

通过这两个边界条件可以得出 $\alpha=20$，$\beta=0$，$S_0=4$，$t_0=2$.

12.6 双变量二阶方程的简单介绍

许多常见图形如圆、椭圆、抛物线、双曲线，它们都可以用二阶方程表示．本节我们将回顾一些几何分析方面的知识，因为它们与偏微分方程中常用术语有关．

首先，令 x，y 表示两个确定性变量．我们可以定义以下二阶方程：

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0 \tag{12.56}$$

这里 A，B，C，D，E，F 表示不同的常数(A，B，C 不同时为 0)．这个方程中 x 和 y 的最高次是二次的，因此它是一个二阶方程．

通过对 A，B，C，D，E，F 取不同的值，我们可以使该方程在 xy-平面上的轨迹成为椭圆、抛物线、双曲线或圆．我们接下来就此做一些简单的介绍．

12.6.1 圆

考虑以下取值：

$$A = C \quad 且 \quad B = 0 \tag{12.57}$$

则原二阶方程将退化为：

$$Ax^2 + Ay^2 + Dx + Ey + F = 0 \tag{12.58}$$

利用完全平方公式，这个方程总可以写作

$$(x - x_0)^2 + (y - y_0)^2 = R^2 \tag{12.59}$$

而这正是圆心为(x_0, y_0)半径为 R 的圆．为了具体了解其原理，将(12.59)展开：

$$x^2 + y^2 - 2x_0 x - 2y_0 y + x_0^2 + y_0^2 = R^2 \tag{12.60}$$

在该方程中，我们可以令

$$A = \frac{1}{R} \tag{12.61}$$

$$\frac{-2x_0}{R} = D \tag{12.62}$$

$$\frac{-2y_0}{R} = E \tag{12.63}$$

$$\frac{x_0^2 + y_0^2}{R} = F \tag{12.64}$$

所以，只要 $A=C$，$B=0$，满足该二阶方程的 x 和 y 在 xy 平面上的轨迹总会是一个圆.

在 $R=0$ 的特殊情况中，圆退化成了一个点. 另一种退化的情况是取 $A=C=0$. 此时圆会退化成一条直线，而方程不存在二阶项.

12.6.2 椭圆

第二种情况是当

$$B^2-4AC<0 \tag{12.65}$$

这与圆的情况类似，但 B 不为零，且 x^2 与 y^2 的系数可能不同. 我们可以将原二阶方程改写为

207

$$\alpha(x-x_0)^2+\beta(y-y_0)^2+\gamma(x-x_0)(y-y_0)=R^2 \tag{12.66}$$

它是一个椭圆方程，中心为(x_0,y_0).

通过比较(12.56)与(12.66)前的系数，我们可以得到含 6 个未知数的 6 个方程，所以只要给出一组 A，B，C，D，E，F 的值，我们总能求出参数 x_0，y_0，α，β，γ，R.

一个简单的例子

利用完全平方公式改写方程的方法可以帮助我们区分椭圆、圆、双曲线和抛物线. 我们用以下简单的例子来说明这一点.[⊖]

考虑以下二阶方程

$$9x^2+16y^2-54x-64y+3455=0 \tag{12.67}$$

注意到

$$B^2-4AC=-576 \tag{12.68}$$

故该方程是椭圆型方程. 我们直接写出完全平方形式：

$$9(x^2-6x+?)+16(y^2-6y+?)=3455 \tag{12.69}$$

在问号出填入适当的数值后，我们就能够使括号内的项构成平方项. 用 9 替换第一个括号，则等式右端要增加 81. 用 4 替换第二个括号，则等式右端要增加 64. 可以得到

$$9(x-3)^2+16(y-2)^2=3600 \tag{12.70}$$

也就是

$$\frac{(x-3)^2}{400}+\frac{(y-2)^2}{225}=1 \tag{12.71}$$

它是中心为 $x=3$，$y=2$ 的椭圆.

12.6.3 抛物线

(12.56)中的二阶方程可以退化为抛物线，只要满足

$$B^2-4AC=0 \tag{12.72}$$

满足该条件最简单的方法是令 $B=0$，且 $A=0$ 或 $C=0$. 当该条件满足时，原二阶方程将退

⊖ 完全平方的方法在几何随机微分方程的计算中非常常用.

化为

$$Ax^2 + Dx + Ey + F = 0 \tag{12.73}$$

它是抛物线方程的一般形式．

12.6.4 双曲线

只要满足条件

$$B^2 - 4AC > 0 \tag{12.74}$$

则(12.56)中的二阶方程是双曲线方程的一般形式．这种类型的方程对我们来说不太常用，因此我们跳过这块内容．

12.7 偏微分方程的类型

例 2 告诉我们 $F(S_t, t)$ 的轮廓一般都是非线性曲线．在例 2 中，轮廓是椭圆．实际上，对于以下形式的偏微分方程：

$$a_0 + a_1 F_t + a_2 F_s + a_3 F_{ss} + a_4 F_{tt} + a_5 F_{st} = 0 \tag{12.75}$$

若满足下列条件，则是椭圆型偏微分方程：

$$a_5^2 - 4a_3 a_4 < 0 \tag{12.76}$$

若满足下列条件，则是抛物型偏微分方程：

$$a_5^2 - 4a_3 a_4 = 0 \tag{12.77}$$

208

若满足下列条件，则是双曲型偏微分方程：

$$a_5^2 - 4a_3 a_4 > 0 \tag{12.78}$$

图 12-3 中的 $F(S_t, t)$ 是一个椭圆型偏微分方程的解的图像，因为 $a_4 = 0$ 且 a_3、a_5 同号．故满足条件

$$a_5^2 - 4a_3 a_4 < 0 \tag{12.79}$$

例子：抛物型偏微分方程

图 12-4 给出了以下 $F(S_t, t)$ 的图像：

$$F(S_t, t) = -10(S_t - 4)^2 - 3(t - 2) \tag{12.80}$$

该函数的轮廓是抛物线．固定 t 时 $F(S_t, t)$ 关于 S_t 为抛物线，固定 S_t 时 $F(S_t, t)$ 关于 t 为直线．

$F(S_t, t)$ 是以下偏微分方程的解：

$$-\frac{1}{4}F_{ss} + \frac{5}{3}F_t = 0 \tag{12.81}$$

由于 $a_4 = 0$ 且 $a_5 = 0$，该方程的系数满足

$$a_5^2 - 4a_3 a_4 = 0 \tag{12.82}$$

故该偏微分方程属于抛物型．

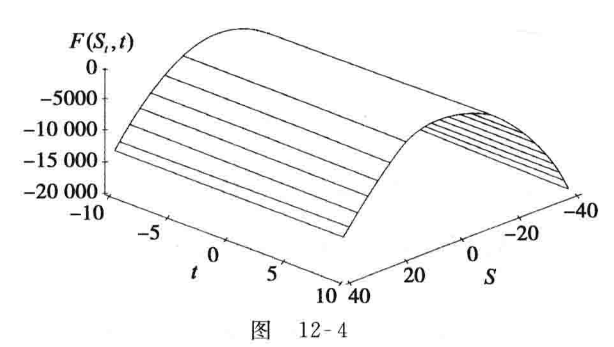


图 12-4

12.8 方差伽马模型定价

我们将根据所研究期权的种类，从分析层面(通过适当转化的方法)和数值层面(通过求解偏微分-积分方程)来介绍 VG 模型下的衍生品定价．对于 VG 模型下的欧式期权定价，首先要以随机时间 g 为条件，然后使用 Black-Scholes 公式来求条件期权价值．所以要求 VG 欧式期权 $C(S_0,K,T)$ 价值，需要关于伽马密度求积分．

209

$$C(S_0,K,T)=\int_0^{\infty}\text{Black-Scholes}(S_0,K,g)\,\frac{g^{t/\upsilon-1}\mathrm{e}^{-g/\upsilon}}{\upsilon^{t/\upsilon}\Gamma(t/\upsilon)}\mathrm{d}g$$

此外，对方程(12.85)和(12.86)应用列维－辛钦定理(Sato，2001)，可知 VG 过程的列维测度可写为 $\mathrm{d}\upsilon(x)=k(x)\mathrm{d}x$，其中 $k(x)$[⊖] 为

$$\mathrm{d}\upsilon(x)=k(x)\mathrm{d}x$$

$$k(x)=\frac{\mathrm{e}^{-\lambda_p x}}{\upsilon x}1_{x>0}+\frac{\mathrm{e}^{-\lambda_n|x|}}{\upsilon|x|}1_{x<0}$$

$$\lambda_p=\left(\frac{\theta^2}{\sigma^4}+\frac{2}{\sigma^2\upsilon}\right)^{\frac{1}{2}}-\frac{\theta}{\sigma^2}$$

$$\lambda_n=\left(\frac{\theta^2}{\sigma^4}+\frac{2}{\sigma^2\upsilon}\right)^{\frac{1}{2}}+\frac{\theta}{\sigma^2}$$

偏微分-积分方程(PIDE)的推导

在我们推导 VG 过程的 PIDE 之前，我们先介绍一下列维过程的一般形式和它的列维-辛钦表达．在金融工程和数理金融中，列维过程是相对比较新的概念，在实践中还没有得到广泛运用．如果读者想进一步了解列维过程和无限可分分布，可以参考 Bertoin(1996)和 Papapantoleon(2008)．它们对于许多概念的理解都有着非常关键的作用，如纯跳跃和一般列维模型、列维过程的分解(分解为布朗运动和泊松跳跃成分)、列维密度的形式的推导．关于跳跃过程在金融建模中的运用，我们可以参考 Cont 和 Tankov(2003)，它对金融中的一般列维模型有很具体的介绍．

列维过程是具有平稳增量的随机过程．列维-辛钦定理通过列维过程特征函数给出了该过程的特征描述：即存在测度 υ 满足对所有 $u\in\mathbf{R}$ 和非负的 t，有

$$E(\mathrm{e}^{iuX_t})=\exp(t\phi(u)) \tag{12.83}$$

其中

$$\phi(u)=i\gamma u-\frac{1}{2}\sigma^2u^2+\int_{-\infty}^{+\infty}(\mathrm{e}^{iuy}-1-iuy1_{\{|y|\leqslant1\}})\mathrm{d}\upsilon(y) \tag{12.84}$$

这里 γ 和 σ 为实数，υ 是 $\mathbf{R}$ 上的测度且满足 $\upsilon(\{0\})=0$ 且 $\int_{-\infty}^{+\infty}\min(1,x^2)\mathrm{d}\upsilon(x)$ 有界．假设一个列维过程 $\{X_t\}_{t\geqslant0}$ 可以写成如下形式：

⊖ $\mathrm{d}\upsilon(x)$ 是这个方差伽马过程在风险中性测度 Q 下的列维测度．

$$X_t = (r - q + \mu)t + Z_t \tag{12.85}$$

该过程包含一个由 μ 控制的漂移项和一个纯跳跃部分．在这里，我们关注的是这样的一类列维过程：它的纯跳跃部分的列维测度可以写成 $d\upsilon(y) = k(y)dy$ 的形式，其中 $k(y)$ 为：

$$\begin{aligned} d\upsilon(y) &= k(y)dy \\ k(y) &= \frac{e^{-\lambda_p y}}{\upsilon y^{1+Y}} 1_{y>0} + \frac{e^{-\lambda_n |y|}}{\upsilon |y|^{1+Y}} 1_{y<0} \\ \lambda_p &= \left(\frac{\theta_2}{\sigma^4} + \frac{2}{\sigma^2 \upsilon}\right)^{\frac{1}{2}} - \frac{\theta}{\sigma^2} \\ \lambda_n &= \left(\frac{\theta_2}{\sigma^4} + \frac{2}{\sigma^2 \upsilon}\right)^{\frac{1}{2}} + \frac{\theta}{\sigma^2} \end{aligned} \tag{12.86}$$

这里变量 Y 是用来控制大小跳跃的符号的．当 Y 大于零时，将在 0 附近活跃度较高，而在远离 0 的位置活跃度较低．此外还有一些有意思的临界值：

- 当 $Y<1$ 时有限变差，当 $Y>1$ 时无限变差，$Y=1$ 是分界点．
- 当 $Y<0$ 时到达概率有限，当 $Y>0$ 时到达概率无限，$Y=0$ 是分界点．
- 当 $Y<-1$ 时远离 0 处活跃度较高，当 $Y>-1$ 时 0 附近活跃度较高，$Y=-1$ 是分界点.

假设函数 $F(S_t, t)$ 代表衍生证券的价值．对 $e^{r(T-t)}F(S_t, t)$ 运用伊藤引理来得到半鞅，可得 210

$$\begin{aligned} F(S_T, T) &= e^{rT}F(S_0, 0) + \int_0^T e^{r(T-t)} \frac{\partial V}{\partial S}(S_t, t)dS_t + \int_0^T e^{r(T-t)} \int_{-\infty}^{+\infty} \\ &\quad \times \left[F(S_{t-}e^x, t) - F(S_{t-}, t) - \frac{\partial V}{\partial S}(S_{t-}, t)S_{t-}(e^x - 1)\right]\mu(dx, dt) \\ &\quad + \int_0^T e^{r(T-t)} \left[\frac{\partial V}{\partial t}(S_t, t) - rF(S_t, t)\right]dt \\ &= e^{rT}F(S_0, 0) + \int_0^T e^{r(T-t)} \frac{\partial V}{\partial S}(S_t, t) \times [dS_t - (r-q)S_t dt] \\ &\quad + \int_0^T e^{r(T-t)} \int_{-\infty}^{+\infty} \times \left[F(S_{t-}e^x, t) - F(S_{t-}, t) - \frac{\partial V}{\partial S}(S_{t-}, t)S_{t-}(e^x - 1)\right]\mu(dx, dt) \\ &\quad + \int_0^T e^{r(T-t)} \left[\frac{\partial V}{\partial t}(S_t, t) + (r-q)S_t \frac{\partial V}{\partial S}(S_t, t) - rF(S_t, t)\right]dt \end{aligned}$$

其中 $u(dx, dt)$ 是整值随机测度，可以对空间一时间范围内的跳跃次数计数．密度 $\upsilon(dy)dt$ 是对 $u(dx, dt)$ 的补充(Jacod 和 Shiryaev，1987)．将以下项添加到上面的方程，可得

$$\int_0^T e^{r(T-t)} \int_{-\infty}^{+\infty} \left[F(S_{t-}e^y, t) - F(S_{t-}, t) - \frac{\partial V}{\partial S}(S_{t-}, t)S_{t-}(e^y - 1)\right]\upsilon(dy)dt$$

$$\begin{aligned} F(S_T, T) &= F(S_0, 0)e^{rT} + \int_0^T e^{r(T-t)} \frac{\partial V}{\partial S}(S_t, t)[dS_t - (r-q)S_t dt] + \int_0^T e^{r(T-t)} \int_{-\infty}^{+\infty} \\ &\quad \times \left[F(S_t e^y, t) - F(S_t, t) - \frac{\partial V}{\partial S}(S_t, t)S_t(e^x - 1)\right] \times [\mu(dx, dt) - \upsilon(dy)dt] \\ &\quad + \int_0^T e^{r(T-t)} \int_{-\infty}^{+\infty} \times \left[F(S_t e^x, t) - F(S_{t-}, t) - \frac{\partial V}{\partial S}(S_t, t)S_{t-}(e^x - 1)\right]\upsilon(du)dt \end{aligned}$$

$$+\int_0^T e^{r(T-t)} \times \left[\frac{\partial V}{\partial t}(S_t,t)+(r-q)S_t\frac{\partial V}{\partial S}(S_t,t)-rF(S_t,t)\right]dt$$

现在，在风险中性测度 Q 下求期望，可得

$$E^Q F(S_T,T)=F(S_0,0)e^{rT}+\int_0^T e^{r(T-t)}\left\{\int_{-\infty}^{+\infty}\left[F(S_t e^y,t)-F(S_t,t)-\frac{\partial V}{\partial S}(S_t,t)S_{t-}(e^y-1)\right]\upsilon(dy)\right.$$
$$\left.+\frac{\partial V}{\partial t}(S_t,t)+(r-q)S_t\frac{\partial V}{\partial S}(S_t,t)-rF(S_t,t)\right\}dt$$

由于

$$E^Q(F(S_T,T))=F(S_0,0)e^{rT}$$

所以

211

$$\int_0^T e^{r(T-t)}\left\{\int_{-\infty}^{+\infty}\left[F(S_{t-}e^y,t)-F(S_{t-},t)-\frac{\partial V}{\partial S}(S_{t-},t)S_{t-}(e^y-1)\right]\upsilon(dy)\right.$$
$$\left.+\frac{\partial V}{\partial t}(S_t,t)+(r-q)S_t\frac{\partial V}{\partial S}(S_t,t)-rF(S_t,t)\right\}dt=0$$

由于被积函数是非负的，那么

$$\int_{-\infty}^{+\infty}\left[F(S_{t-}e^y,t)-F(S_{t-},t)-\frac{\partial V}{\partial S}(S_{t-},t)S_{t-}(e^y-1)\right]\upsilon(dy)$$
$$+\frac{\partial V}{\partial t}(S_t,t)+(r-q)S_t\frac{\partial V}{\partial S}(S_t,t)-rF(S_t,t)=0$$

注意该 PIDE 对于任何列维密度 $\upsilon(dy)$ 都具有一般性．由 $\upsilon(dy)=k(y)dy$ 我们可得

$$\int_{-\infty}^{\infty}\left[F(S_{t-}e^y,t)-F(S_{t-},t)-\frac{\partial V}{\partial S}(S_{t-},t)S_{t-}(e^y-1)\right]k(y)dy$$
$$+\frac{\partial V}{\partial t}(S_t,t)+(r-q)S_t\frac{\partial V}{\partial S}(S_t,t)-rF(S_t,t)=0 \tag{12.87}$$

它是偏微分-积分方程．

12.9 小结

本章我们介绍了偏微分方程的概念．这类方程的解是所求变量的函数．我们简单地介绍了偏微分方程的各种形式及相关术语．

本章我们还展示了如何利用金融衍生品和标的资产的关系得到偏微分方程，由此得到的偏微分方程是衍生资产价格必须满足的．

12.10 参考阅读

由于偏微分方程在衍生资产定价的实践中非常有用，大多数读者都会对其感兴趣．因此，除了本书介绍的偏微分方程理论之外，偏微分方程数值解方面的书籍也非常有用．通常这类书籍也包含了对相应理论的简单介绍．在这里向我们向有兴趣的读者推荐两本书，Smith (1985)和 Thomas(1995)，前者的内容更容易理解，而后者在内容上更全面，也距今更近．

12.11 习题

1. 已知含3个变量 x，y，z 的方程 $f(x,y,z)$，以下偏微分方程叫做拉普拉斯方程：

$$f_{xx}+f_{yy}+f_{zz}=0$$

可以看到，拉普拉斯方程中，关于3个变量各自的二阶偏导和等于0.

(a) 以下函数是否满足拉普拉斯方程？

i) $f(x,\ y,\ z)=4z^2y-x^2y-y^3$

ii) $f(x,\ y)=x^2-y^2$

iii) $f(x,\ y)=x^3-3xy$

iv) $f(x,\ y,\ z)=\dfrac{x}{y+z}$

(b) 为什么存在不止一个函数满足拉普拉斯方程？一个方程一般会有许多解，这是一件好事吗？

2. 四变量函数 $f(x,\ y,\ z,\ t)$满足以下热传导方程：

$$f_t=a^2(f_{xx}+f_{yy}+f_{zz})$$

其中 a 是常数．根据热传导方程，$f(\cdot)$关于 t 的一阶偏导与 $f(\cdot)$关于其余变量的二阶偏导之和成比例. 以下函数是否满足热传导方程？ 212

(a) $f(x,y,z)=\mathrm{e}^{[29a^2\pi^2t+\pi(3x+2y+4z)]}$

(b) $f(x,y,z)=3x^2+3y^2-6z^2+x+y-9z-3$

3. 考虑偏微分方程：

$$f_x+0.2f_y=0$$

其中 $x\in[0,1]$且 $y\in[0,1]$.

(a) 该方程中的未知量是什么类型的？

(b) 用简单直接的语言解释该方程的含义．

(c) 你能找到多少个满足该方程的函数 $f(x,y)$？

(d) 现假设已知边界条件：$f(0,Y)=1$. 你能否找到一个偏微分方程的解？该解是否唯一？

4. 考虑以下偏微分方程：

$$f_{xx}+0.2f_t=0$$

边界条件为

$$f(x,1)=\max[x-6,0]$$

且 $0\leqslant x\leqslant 12$，$0\leqslant t\leqslant 1$

(a) 仅凭这一个边界条件，能否得到 $f(x,t)$的数值近似？

(b) 添加一个边界条件，你可选择 $f(0,t)$或 $f(12,t)$两个位置．

(c) 根据你添加的边界条件，取网格间隔为 $\Delta x=3$ 和 $\Delta t=0.25$，计算 $f(x,y)$的数值近似．

5. 用偏微分方程考虑以下期权定价问题：

(a) 欧式看涨期权

(b) 欧式看跌期权

(c) 上升敲出看涨期权

(d) 下降敲出看跌期权

对于每种情况，给出适当的关于 $f(S_{\min},t)$或 $f(S_{\max},t)$或到期日 T 的边界条件，使得我们可以在求解偏微分方程时使用．

6. 写一个能为欧式定点期权定价的拟合程序，它的标的股票的价格过程是一个几何布朗运动，参数为 $\sigma=0.1$，$r=0.05$，$q=0$，$S=K=10$，$T=1$. 213

第13章 偏微分方程与偏积分-微分方程——一个应用

13.1 引言

本章我们将介绍一些用偏微分方程法对衍生资产定价的例子.

我们这样做的目的，是从几何的角度观察由 Black 和 Scholes(1973)得到的偏微分方程的解函数. Black-Scholes 公式的几何意义能够帮助我们理解偏微分方程. 特别地，我们将介绍单随机变量看涨期权的几何含义.

215

然后，我们将引入第二个因子，将 Black-Scholes 的原始结构复杂化. 这会造成一些难题，我们将就这一点简要介绍.

最后，我们将比较偏微分方程的解析解与数值解，并以一个资产价格数值计算的例子作为本章的结束.

13.2 Black-Scholes 偏微分方程

在第12章中，我们介绍了关于标的资产 S_t 的衍生品所必需满足的条件. 标的证券不支付红利，且假设无风险利率为常数 r.

现在，假设我们所考虑的 SDE 中

$$a(S_t,t)=\mu S_t \tag{13.1}$$

$$\sigma(S_t,t)=\sigma S_t \quad t\in[0,\infty) \tag{13.2}$$

我们有时候用 σ_t 代替 σS_t. 在这些条件下，Black-Scholes 基础偏微分方程以及边界条件为：

$$-rF+rF_tS_t+F_t+\frac{1}{2}\sigma^2F_{ss}S_t^2=0,\ 0\leqslant S_t,0\leqslant t\leqslant T \tag{13.3}$$

$$F(T)=\max[S_T-K,0] \tag{13.4}$$

方程(13.3)和(13.4)在1973年首次被 Black 和 Scholes 应用在金融中. 因此我们称该方程为“Black-Scholes 基础偏微分方程”. [⊖]

Black 和 Scholes 求解了该偏微分方程，并得到了解函数 $F(S_t,t)$的确切表达式：

$$F(S_t,t)=S_tN(d_1)+K\mathrm{e}^{-r(T-t)}N(d_2) \tag{13.6}$$

其中

⊖ 该 PDE 中只出现了一个二阶偏导，即关于 S_t 的偏导. 同时，方程中不含有常数项. 在这些条件下，我们很容易发现第12章中所介绍的表达式(13.5)的值总为0：

$$a_5^2-4a_3a_4 \tag{13.5}$$

这意味着不管 S_t 是多少，Black-Scholes 方程总是抛物线型的.

$$d_1 = \frac{\log\left(\frac{S_t}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}} \tag{13.7}$$

$$d_2 = d_1 - \sigma\sqrt{T-t} \tag{13.8}$$

$N(d_i)(i=1,2)$是两个标准正态密度函数的积分：

$$N(d_i) = \int_{-\infty}^{d_i} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}x^2} \mathrm{d}x \tag{13.9}$$

要证明该函数满足 Black-Scholes 偏微分方程及边界条件，我们要对(13.6)关于 S_t 和 t 求一阶或二阶偏导，然后按照(13.3)组合起来，得到的和应该为 0，而当 t 趋向于 T 时，函数应该满足(13.4).

从几何角度观察 Black-Scholes 公式

在第 12 章中我们看到了一些偏微分方程的解在三维空间中的图像，在这里，对于 Black-Scholes 偏微分方程，我们也可以类似地将其解函数在三维空间中表达出来．本偏微分方程的解已由(13.6)给出．我们希望对参数 K，r，σ，T 取适当的值，在三维空间 $F\times s\times t$ 中画出该解．

我们取

$$r = 0.065, \quad K = 100, \quad \sigma = 0.080, \quad T = 1 \tag{13.10}$$

将这些值代入(13.6)．这些值意味着，无风险利率为 6.5%，在 $t\in[0,1]$期间波动率为 80%．这样的波动率对于大多数成熟的金融市场是偏高的．但这样的取值可以让图像特征更明显．这个看涨期权的期限被标准化为 1，$T=1$ 表示一年，且初始时间设为 $t_0=0$．执行价格设为 100.⊖

216

为根据这些取值画出 Black-Scholes 公式的图像，我们需要先确定自变量 S_t 和 t 的取值范围．在这里，我们取 S_t 的范围为 50 到 140，t 为 0 到 1，所得到的曲面如图 13-1 所示．

在图 13-1 中，有一个 S_t 轴和 T-t 围成的“水平面”，其中 $T-t$ 代表当前时刻距离到期日的时间．这两条轴线确定了一个平面．例如，点 A 表示标的资产价格为 130，距离到期日时间为 0.8．垂直于平面向上作垂线可得点 B，即使点 A 处 Black-Scholes 公式的值：

$$B(130, 0.2) \tag{13.11}$$

图 13-1 中还展示了曲面上两条通过不同方式得到的曲线．首先，固定 S_t，改变 t 的值，我们可以得到类似于图中 aa'的曲线．这里，aa'表示当 S_t 固定为 100、t 从 0 增加到 1 的过程中看涨期权价格的变化情况．

⊖ 如果 $T=1$ 表示“一年”，那么利率和波动率也是按年计算的．但 $T=1$ 也可以表示 6 个月，3 个月，或者任何对金融工具适用的时间段．我们仅把 $T=1$ 当作一种标准化．在这样的条件下，利率和波动率也必须根据相应的时间段调整．

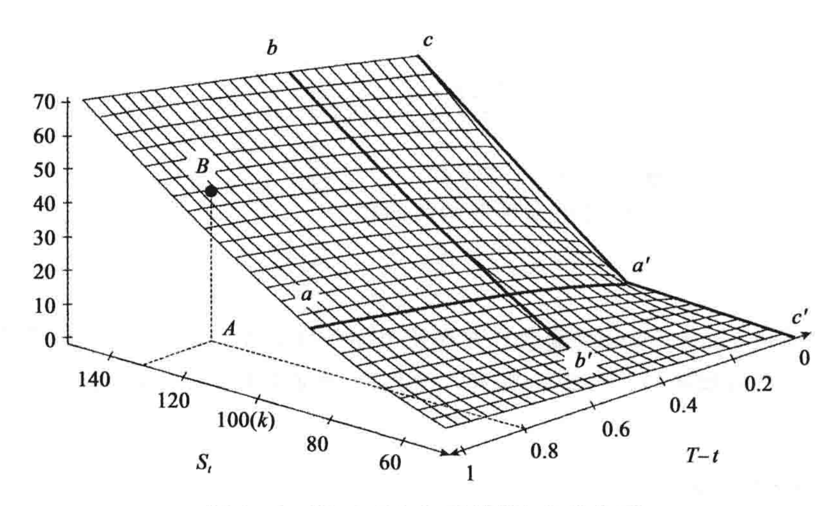


图 13-1 Black-Schole 看涨期权价格曲面

第二种曲线是形如图中 bb'这样的曲线．这里，bb'表示当 t 固定为 0.4，S_t 从 60 增加到 140 的过程中 $F(S_t,t)$的变化情况．有意思的是当 t 趋于 1 时，轮廓趋于 cc'．cc'在 K 处一般会有一个角度，K 是期权到期日的执行价格．

在使用图 13-1 观察衍生品的定价时，读者可能会疑惑如何用一个确定性的曲面去体现随机的衍生品价格过程．一旦我们固定 K，r 和 σ，Black-Scholes 公式得到的是一个曲面. 即使新的随机事件发生、dW_t 变为已知，这个曲面本身是不会移动的．维纳增量的实现只会造成衍生品价格在这个曲面上的随机移动．在本章后文中的图 13-6 中给出了一个这样的例子：图中轨迹 C_0C_T 标记出了衍生品价格在这个曲面上移动过程．由于维纳过程的增量是不可预测的，股票价格沿 t 方向的移动每一步都是随机的．在时间间隔无穷小的情况下，每一步的位移也是无穷小，也还是不可预测的．

轨迹 C_0C_T 还有一个有意思的地方．随着时间推移，S_t 在 $S\times t$ 平面上的轨迹形成一条曲线．而这条轨迹曲线在 B-S 函数曲面上的投影就是轨迹 C_0C_T. 这两条轨迹有着确定性的对应关系．换句话说，已知水平面上 S_t 的轨迹，就可以唯一地确定曲面上 $F(S_t,t)$的轨迹. 这是因为 S_t 与 $F(S_t,t)$具有相同的随机成分．

13.3 局部波动率模型

下面我们考虑另一个特殊的 SDE. 在这个 SDE 中，我们假设漂移项与之前的 Black-Scholes 模型一样：

$$a(S_t,t)=\mu S_t \tag{13.12}$$

但同时，我们对扩散项做如下假设：

$$\sigma(S_t,t)=\sigma(S_t,t)S_t,\quad t\in[0,\infty) \tag{13.13}$$

可以看到，与 Black-Scholes 中 σ 取常数不同，这里它与价格和时间都有关．我们称这样的模型为局部波动率模型．这个模型的基础偏微分方程及其对应的边界条件为

$$-rF+rF_tS_t+F_t+\frac{1}{2}\sigma(S_t,t)^2F_{ss}S_t^2=0,\quad 0\leqslant S_t,\quad 0\leqslant t\leqslant T \tag{13.14}$$

$$F(T)=\max[S_T-K,0] \tag{13.15}$$

与 Black-Scholes 模型不同，对于局部波动率模型，我们找不到其对应偏微分方程的解析解．因此，大部分情况下我们只能利用一些数值方法来得到近似解．

217

由于波动率与股票价格和时间有关，局部波动率模型可以更准确地描述偏度和峰度，这在一定程度上弥补了 Black-Scholes 模型的不足．

13.4 偏微分-积分方程

改写 Black-Scholes 偏微分方程

在给出正式的介绍前，我们先从 Black-Scholes 模型出发，写出它的偏积分-微分方程，并探究当我们研究纯跳跃过程时会发生什么．这是一个反向工程问题．

考虑 Black-Scholes 偏微分方程

$$\frac{\partial F}{\partial t}+\frac{\sigma^2S^2}{2}\frac{\partial^2F}{\partial S^2}+(r-q)S\frac{\partial F}{\partial S}=rF(S,t) \tag{13.16}$$

边界条件的选取和我们所研究期权的具体种类有关，在这里我们先不考虑边界条件．首先我们做变量代换，令 $x=\ln S$ 且 $\widetilde{F}(x,t)=F(S,t)$；注意到根据链式法则我们有

$$\frac{\partial\widetilde{F}}{\partial x}(x,t)=S\frac{\partial F}{\partial S}(S,t) \tag{13.17}$$

$$\frac{\partial^2\widetilde{F}}{\partial x^2}(x,t)-\frac{\partial\widetilde{F}}{\partial x}(x,t)=S^2\frac{\partial^2F}{\partial S^2}(S,t) \tag{13.18}$$

利用以上两式，可将(13.16)改写为

$$\frac{\partial\widetilde{F}}{\partial t}+(r-q)\frac{\partial\widetilde{F}}{\partial x}+\frac{\sigma^2}{2}\left(\frac{\partial^2\widetilde{F}}{\partial x^2}-\frac{\partial\widetilde{F}}{\partial x}\right)=r\widetilde{F}(S,t) \tag{13.19}$$

我们先不合并同类项，后面会看到为什么要这样做．如果 $k(y)$定义为

$$k(y)=\begin{cases}\dfrac{e^{-\lambda|y|}}{\upsilon|y|} & y<0\\[2ex] \dfrac{e^{-\lambda y}}{\upsilon y} & y>0\end{cases}$$

218

其中 $\lambda=\sqrt{\frac{2}{\sigma^2\upsilon}}$. 当 υ 趋于 0 时，我们有

$$\int_{|y|\leqslant\epsilon}y^2k(y)\mathrm{d}y=\sigma^2 \tag{13.20}$$

这里有一个问题：列维密度函数 $k(y)$从哪里来的？参数 υ 和 λ 应该如何取值？它是否与 ϵ 独立？如果知道了这些，我们就可以将(13.19)改写为

$$\frac{\partial\widetilde{F}}{\partial t}+(r-q)\frac{\partial\widetilde{F}}{\partial x}+\int_{|y|\leqslant\epsilon}\frac{1}{2}\left(\frac{\partial^2\widetilde{F}}{\partial x^2}-\frac{\partial\widetilde{F}}{\partial x}\right)\times y^2k(y)\mathrm{d}y=r\widetilde{F}(x,t) \tag{13.21}$$

之所以我们可以把求偏导与求积分的次序调换，因为它们是关于 y 独立的．现在我们想要改写(13.21)中的被积函数．我们考虑类似于 $\widetilde{F}(x+y,t)$的项，其中 y 是跳跃幅度．和 Merton 跳跃-扩散模型一样，我们考虑使用(e^y-1)．对 $\widetilde{F}(x+y,t)$使用泰勒展开，可得

$$\widetilde{F}(x+y,t)=\widetilde{F}(x)+y\frac{\partial\widetilde{F}}{\partial x}+\frac{y^2}{2}\frac{\partial^2\widetilde{F}}{\partial x^2}+O(y^3) \tag{13.22}$$

对 e^y 使用泰勒展开，得

$$e^y=1+y+\frac{y^2}{2}+O(y^3) \tag{13.23}$$

在以上两个展开式中，余项都是 y^3 阶的无穷小量．根据(13.22)我们有

$$\frac{y^2}{2}\frac{\partial^2\widetilde{F}}{\partial x^2}=\widetilde{F}(x+y,t)-\widetilde{F}(x)-y\frac{\partial\widetilde{F}}{\partial x}+O(y^3) \tag{13.24}$$

根据(13.23)我们有

$$\frac{y^2}{2}=e^y-1-y+O(y^3) \tag{13.25}$$

将以上两式带入(13.21)，可以得到

$$\frac{\partial\widetilde{F}}{\partial t}+(r-q)\frac{\partial\widetilde{F}}{\partial x}+\int_{|y|\leqslant\varepsilon}\left(\widetilde{F}(x+y)-\widetilde{F}(x)-y\frac{\partial\widetilde{F}}{\partial x}-(e^y-1-y)\frac{\partial\widetilde{F}}{\partial x}\right)k(y)\mathrm{d}y=r\widetilde{F}(x,t) \tag{13.26}$$

或者，进一步化简可得

$$\frac{\partial\widetilde{F}}{\partial t}+(r-q)\frac{\partial\widetilde{F}}{\partial x}+\int_{|y|\leqslant\varepsilon}\left(\widetilde{F}(x+y)-\widetilde{F}(x)-(e^y-1)\frac{\partial\widetilde{F}}{\partial x}\right)k(y)\mathrm{d}y=r\widetilde{F}(x,t) \tag{13.27}$$

这里我们省略了余项$(O(y^3))$．因为已知$|y|\leqslant\varepsilon$ 且假设 ε 非常小，$(O(y^3))$项可以忽略不计．由于 $S=e^x$ 和 $\widetilde{F}(x+y,t)=F(Se^y,t)$，将其代回上式，可得

$$\frac{\partial F}{\partial t}+(r-q)S\frac{\partial F}{\partial S}+\int_{|y|\leqslant\varepsilon}\left(F(Se^y,t)-F(S,t)-(e^y-1)\frac{\partial F(S,t)}{\partial S}\right)k(y)\mathrm{d}y=rF(S,t) \tag{13.28}$$

关于以上推导，我们列出了一些需要注意的地方．

- 对于 Black-Scholes 模型这样的纯扩散模型而言，我们可能得到一个良态的扩散项．但是对于这节中我们讨论的带跳模型而言，其二阶偏导可能不存在．因此我们用一个积分区间很小的积分代替扩散项，而这个积分中的被积函数不含 F 的二阶偏导数.
- 函数 $k(y)$有着类似于计数的功能，它记录了跳跃的次数．而该问题中有无穷多次的小跳跃．
- 从这里我们可以猜测，如果要描述一个纯跳跃过程，不管这个过程中发生大幅度跳跃的频率有多低，我们都还是需要一个同样适用于大幅度跳跃的积分来代替纯扩散模型中的扩散项．

- 我们将这视为一个良好的开端．我们知道了在某些特殊情况下(比如市场是满足 Black-Scholes 条件)，Black-Scholes 方程中的扩散项可以被写成一个积分——这样纯跳跃过程也可以看作为一个 Black-Scholes 方程的特例．

219

- 如何用积分一微分方程表述纯跳跃过程．
- 如何贴切地描述纯跳跃过程．
- $k(y)$是什么，怎么解释 $k(y)$，如何得到它．
- Black-Scholes 与纯跳跃过程有关联，并且能够描述它，这一点是非常吸引人．
- 模型的结果最好能够符合波动率的微笑．
- 存在一个相容的理论框架．

13.5 资产定价中的偏微分方程/偏积分-微分方程

Black 和 Scholes 是在一些具体的假设条件下得到他们所研究的偏微分方程的．这些基本假设是(1)标的资产是一只股票．(2)该股票不支付股息．(3)该衍生资产是欧式看涨期权，不能在到期日之前执行．(4)无风险利率是常数．(5)不考虑资产的不可分割性以及可能产生的交易费用，如佣金差价或者买卖差价．

在大多数实际问题中，以上假设中的一个或多个可能不成立，这样，Black-Scholes 偏微分方程将不再适用，而我们因此需要找出另一个新的偏微分方程．但有一个例外：如果只调整假设(3)为美式看涨期权，那得到的偏微分方程将不会有任何变化．

这些调整后更加复杂的偏微分方程，有一些也得到了广泛的运用．我们接下来将介绍一个简单的例子．

常数股息

如果我们想对某看涨期权定价，而该期权的标的股票支付股息，单位时间的股息率为 σ，那么对应的偏微分方程将略微调整．

假定我们调整 Black-Scholes 假设，引入股息率为常数 σ 的标的资产 S_t.

我们还是能够利用标的资产和它的看涨期权，建立“近似的”无风险投资组合：

$$P_t = \theta_1 F(S_t, t) + \theta_2 S_t \tag{13.29}$$

组合权重 θ_1、θ_2 取为：

$$\theta_1 = 1, \quad \theta_2 = - F_s \tag{13.30}$$

因此不可预测的随机成分被消除了，该组合能够得到稳定的收益：

$$\mathrm{d}P_t = F_t \mathrm{d}t + \frac{1}{2} F_{ss} \sigma_t^2 \mathrm{d}t \tag{13.31}$$

截至这一步，与第 12 章中的原始 Black-Scholes 方法没有任何差异．P_t 的时间路径也是完全可预测的．

区别在于对资产组合升值率的确定．之前，(完全可预测的)资产收益恰好等于无风险投资收益．但现在，标的股票按比率 σ 支付股利．所以，无风险投资组合的收益等于资产

收益加上收到的股利：

$$dP_t + \sigma dt = rP_t dt \tag{13.32}$$

或者

$$dP_t = -\sigma dt + rP_t dt \tag{13.33}$$

将它与(13.31)结合起来，可以得到一个与(13.14)略微不同的偏微分方程．

$$rF - rF_s S_t - \sigma - F_t - \frac{1}{2} F_{ss} \sigma_t^2 = 0 \tag{13.34}$$

现在出现了一个常数项 σ．所以可见，股票按常利率 σ 支付股利并没有对问题产生很大困难．

220

13.6 奇异期权

上一节我们讨论了一种基于 Black-Scholes 框架的复杂情况．随着关于股利支付的假设改变，衍生资产无套利价格的偏微分方程也发生了改变．本节将讨论另一种特殊情况．

假设所研究的衍生资产是期权，而到期日是随机的．例如，“下降敲出”“上升敲出”等障碍期权．[1]与标准期权不同，这类工具的最终支付取决于期限内标的资产价格是否达到某个障碍．如果达到，期权的价值会改变．我们简单地介绍几种“奇异”期权．

13.6.1 回望期权

在标准 Black-Scholes 模型中，如果到期日期权是实值的，看涨期权的价值为 $S_T - K$．其中 S_T 是标的资产在到期日的价格，常数 K 是执行价格．

在浮动回望期权中，到期日的实值看涨期权价值是 $S_T - S_{\min}$，其中 $S_{\min}$ 是标的资产在期权期限内出现的最低价格．

而对于固定回望看涨期权，到期日价值为固定的执行价格 K 与 $S_{\max}$ 之差(如果为正的话)．$S_{\max}$ 是标的资产在期权期限内出现的最高价格．[2]只要期权在其期限内的任意一段时间是实值的，那么该期权就能够获得正收益．所以，如果其他条件一致，这种期权比标准看涨期权的价格更高．

13.6.2 梯式期权

梯式期权包含若干个梯级价格．如果标的价格到达这些梯级价格，一部分期权的回报就会被“锁定”．

13.6.3 触发式或敲入期权

对于下降-敲入期权，在期权期限内，当标的资产现货价格下降到某障碍以下，持有者将

[1] 它们也叫“敲出”和“敲入”期权．

[2] 回望期权是浮动的，因为执行价格没有固定．

得到一个欧式期权．如果始终没有跌破该障碍，期权过期时持有者会获得少量补偿．[1]

13.6.4 敲出期权

敲出期权是一种特殊的欧式期权，当标的资产价格下降到某障碍以下，该期权会立即失效．如果到达了该障碍，期权会给予少量补偿．如果没有到达，该期权等价于标准欧式期权．这样的期权叫做"下降-敲出"期权．[2]

13.6.5 其他奇异期权

我们可以用许许多多不同的方式构建特种期权．一些常见的情形包括：

- 篮子期权．其标的资产是多种不同的金融工具的组合．这样的投资组合可以弱化单个证券的波动率．在新兴市场的金融衍生品中，篮子期权已经变得越来越被人们接受．
- 多资产期权的价值依赖于多个标的资产的价格．例如，在到期日这种看涨期权的价值可能为

221

$$F(S_{1T},S_{2T},T) = \max[0,\max(S_{1T},S_{2T}) - K] \tag{13.35}$$

以及价差看涨期权

$$F(S_{1T},S_{2T},T) = \max[0,(S_{1T} - S_{2T}) - K] \tag{13.36}$$

或者组合看涨期权

$$F(S_{1T},S_{2T},T) = \max[0,(\theta_1 S_{1T} + \theta_2 S_{2T}) - K] \tag{13.37}$$

其中 θ_1、θ_2 是组合权重．最后再举一个例子，还有一种双重执行看涨期权：

$$F(S_{1T},S_{2T},T) = \max[0,S_{1T} - K,S_{2T} - K] \tag{13.38}$$

- 平均期权，也称为亚式期权，也是常见期权．其价格取决于标的资产在期权期限内的平均价格．[3]

13.6.6 奇异期权的偏微分方程

从以上奇异期权的简单举例中可以看出，奇异期权与 Black-Scholes 模型的标准情况主要有 3 个区别．

第一，期权在到期日的价值可能受期限内发生的特殊事件的影响(如标的资产在期限内最高价格的函数)．显然，这使得边界条件比标准 Black-Scholes 模型更加复杂．

第二，衍生工具的到期日可能是随机的．

第三，衍生工具可能建立在超过一种资产上．这些区别都会改变原模型中的基础偏微分方程．

---

[1] 类似地，存在上升敲入期权，这种期权在标的资产价格上升到一定水平后开始生效．

[2] 上升-敲出期权在标的资产价格上升到某水平后立即失效．

[3] 通常会使用算术平均，用计算机以天、周或月为单位计算．

这里无法介绍所有例子，仅考虑敲出期权的情况. 下面我们讨论下降一敲出看涨期权. 用 K_t 表示 t 时刻的障碍价格，S_t 表示标的资产价格，$F(S_t,t,K_t)$表示敲出期权价格. 如果在期限内，S_t 达到了 K_t，期权持有者将得到 R_t 的补偿，而期权立即失效. 否则，它仍是标准欧式期权.

这样的期权所对应的偏微分方程，与标准 Black-Scholes 模型主要的差别在于边界条件. 当标的资产价格维持在障碍价格 K_t 之上时($t\in[0,T]$)，偏微分方程与标准模型相同：

$$\frac{1}{2}\sigma_t^2 F_{ss} + rF_s S_t - rF + F_t = 0, \quad 若\ S_t > K_t \tag{13.39}$$

且

$$F(S_T,T,K_T) = \max[0, S_T - K_T] \tag{13.40}$$

但是，在期权期限内，只要 S_t 下降到 K_t 以下，那么

$$F(S_t,t,K_t) = R, \quad 若\ S_t \leqslant K_t \tag{13.41}$$

所以偏微分方程的形式是一样的，只是边界条件不同. 这会使得方程的解 $F(S_t,t,K_t)$不同.

13.7 实际中求解偏微分方程/偏积分-微分方程

在实践中，一旦交易者得到了描述衍生品价格 $F(S_t,t)$变化行为的偏微分方程，有两种方法能够计算衍生品的价值.

13.7.1 封闭形式的解

第一种方法与 Black-Scholes 方法相同，涉及求封闭形式的解. 但是描述衍生品价格行为的偏微分方程并不是都能够求出封闭形式的解. 大部分情况下，这类偏微分方程要么非常难解，要么就不存在封闭形式的解.

222

首先，我们介绍一下偏微分方程的封闭形式解和数值解的区别. 函数是某偏微分方程的解，要求函数的偏导数满足方程，如

$$\frac{1}{2}\sigma_t^2 F_{ss} + rF_s S_t - rF + F_t = 0, \quad 0 \leqslant S_t, \quad 0 \leqslant t \leqslant T \tag{13.42}$$

可能存在这种情况：我们可以找到一个连续平面，其偏导数满足偏微分方程. 但该平面不能像 Black-Scholes 模型那样用公式便捷地表示出来. 也就是说，尽管解存在，这个解不能用 S_t 和 t 的函数直接表示.

我们用一个类似的例子介绍这一点. 考虑图 13-2 中的函数 $F(t)$.

从图上可以看出，$F(t)$是连续且光滑的. 所以，在图示的范围内，$F(t)$具有关于时间的导数. 但是，$F(t)$是随手画出的，没有理由认为它恰好能够表示成若干个含 t 的项组成的函数. 例如，指数型函数

$$F(t) = a_2 e^{a_1 t} + a_3 \tag{13.43}$$

其中 $a_i(i=1, 2, 3)$为常数，该函数不能够表示该曲线. 实际上，对于大多数连续光滑函

数，都不能用封闭形式的公式表达.㊀㊁

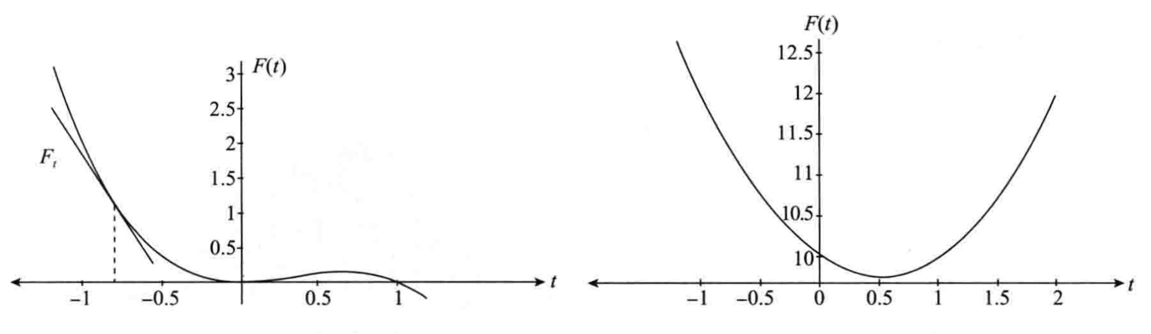


图 13-2 一个无封闭表达式的光滑函数　　图 13-3 一个具有简单封闭表达式的函数

Black-Scholes 模型中偏微分方程的解是一个三维空间中的曲面，坐标轴分别为 S_t，t 和$F(S_t,t)$. 对于三维空间 $F\times t\times S_t$ 上的一个曲面，就算曲面上任意一条曲线都是连续光滑且偏导数存在、满足某个偏微分方程的，这个曲面也可能无法用简洁的公式表示出来.

所以，有可能偏微分方程的解是存在的，但是不能用封闭形式的公式表达出来. 事实上，三维(或更高维的)空间中，用解析式来表示光滑曲面的局限性非常大，大多数情况下这些曲面不能用公式直接表示出来.

13.7.2 数值解

如果封闭形式的解不存在，市场参与者需要去寻找偏微分方程的数值解. 数值解就是直接计算 $F(S_t,t)$曲面上某些点的具体数值，而不是先去求 $F(S_t,t)$的封闭公式. 我们还是考虑 Black-Scholes 框架中的偏微分方程：

$$\frac{1}{2}\sigma_t^2F_{ss}+rF_sS_t-rF+F_t=0,\quad 0\leqslant S_t,\quad 0\leqslant t\leqslant T \tag{13.44}$$

要求出该偏微分方程的数值解，我们要假定该偏微分方程对于 S_t 和 t 的有限增量是成立的. 首先，我们需要对这两个变量进行分割.

1) S_t 网格宽度 ΔS_t 可取为标的证券价格的最小变化单位.

2) 时间 t 是函数 $F(S_t,t)$的另一个变量. 所以，需要取网格宽度 Δt. 不用说，ΔS_t、Δt 必须非常小. 到底需要取到多小，我们需要通过反复试验来确定.

3) 接下来，我们要确定 S_t 的可能取值范围. 为了更准确，我们首先要选取合适的最小值 $S_{\min}$和最大值 $S_{\max}$. 而 S_t 的其他观察值均满足

$$S_{\min}\leqslant S_t\leqslant S_{\max} \tag{13.45}$$

223

4)必须确定边界条件.

㊀ 另一方面，如果曲线的形式比较"特殊"，我们可以进行辨认，并用简单的多项式加以表示. 例如，图 13-3 中的曲线看起来像抛物线，我们就可以用简单的封闭式表达公式 $F=a_0+a_1t+a_2t^2$ 来描述.

㊁ 如果曲线光滑且连续，就可能写成无穷项的泰勒展开式. 但泰勒展开式并不是封闭形式的表达式.

5）假定对于有限小的 ΔS_t 和 Δt 该偏微分方程都是成立的，函数 $F(S_t,t)$ 在网格点上的近似值可以确定下来．

为了解释最后一点，令

$$F_{ij} = F(S_i, t_j) \tag{13.46}$$

其中 F_{ij} 是当时间为 t_j 标的资产价格为 S_i 时的价值，而 i 和 j 的取值范围取决于 ΔS、Δt 和 $S_{\min}$、$S_{\max}$ 的选择．

我们想用有限个点 F_{ij} 来近似 $F(S_t,t)$. 图 13-4中我们用这种方法近似一个任意曲面，图 13-5 中则是近似 Black-Scholes 曲面．在这两幅图中，圆点处代表 $F(S_t,t)$ 被估值的点. ΔS_t 和 Δt 的大小决定了这些点的间距有多近．这些点越近，近似程度越高．

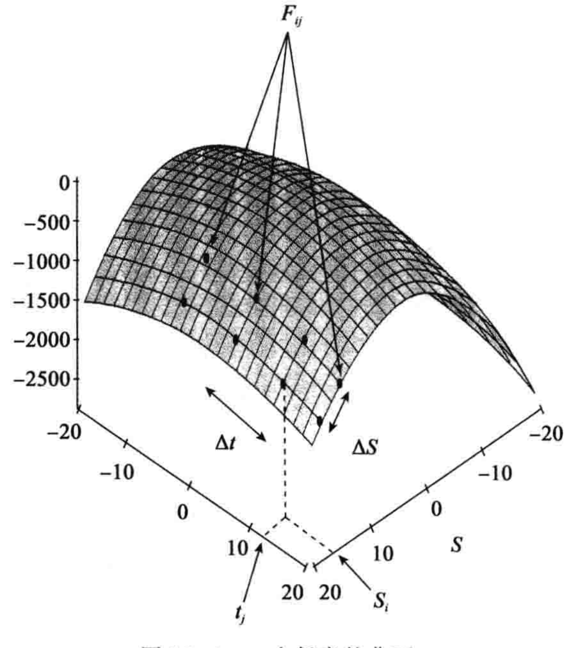


图 13-4　一个任意的曲面

圆点 F_{ij} 对应的位置是 S_t 的第 i 个值和 t 的第 j 个值．首先在坐标轴上选好 S_t 和 t 的值，再计算出该位置 $F(S_t,t)$ 的函数值，将结果记为 F_{ij}.

为进行近似计算，我们需要用合适的 [224] 差分将所有微分替换掉，从而将偏微分方程转化为差分方程．有很多种方法来进行这种差分近似，而在不同方法下最后数值解的精度会有所不同．这里，我们使用最简单的方法㊀

$$\frac{\Delta F}{\Delta t} + rS\,\frac{\Delta F}{\Delta S} + \frac{1}{2}\sigma^2 S^2\,\frac{\Delta^2 F}{\Delta S^2} \approx rF \tag{13.47}$$

其中一阶偏导数用对应的差分来近似．对于一阶偏导我们使用向后差分：

$$\frac{\Delta F}{\Delta t} \approx \frac{F_{ij} - F_{i,j-1}}{\Delta t} \tag{13.48}$$

$$\frac{\Delta F}{\Delta S} \approx \frac{F_{ij} - F_{i-1,j}}{\Delta S} \tag{13.49}$$

或者，我们也可以使用向前差分㊁

$$rS\,\frac{\Delta F}{\Delta S} \approx rS_j\,\frac{F_{i+1,j} - F_{i,j}}{\Delta S} \tag{13.50}$$

对于二阶偏导，我们所用的近似值为

$$\frac{\Delta^2 F}{\Delta S^2} \approx \left[\frac{F_{i+1,j} - F_{i,j}}{\Delta S} - \frac{F_{i,j} - F_{i-1,j}}{\Delta S}\right]\frac{1}{\Delta S} \tag{13.51}$$

㊀ 为方便起见，我们省去下标 i，j. 下面我们会看到，不同方程中元素的 i，j 不同，相应的值也不同．对于每个 i，j，都存在一个(13.47)那样的方程．

㊁ 我们也可以使用中心差分．

其中 $i=1, \cdots, n$ 且 $j=1, \cdots, N$. 参数 N 和 n 决定了曲面 $F(S_t, t)$ 上我们需要计算的点的数量.

例如，在图 13-5 中我们令 $n=5$，$N=22$. 那么，除去边界上的点，在曲面上我们总共需要计算 80 个点的值. 这些值可以通过求解方程组(13.47)得到.

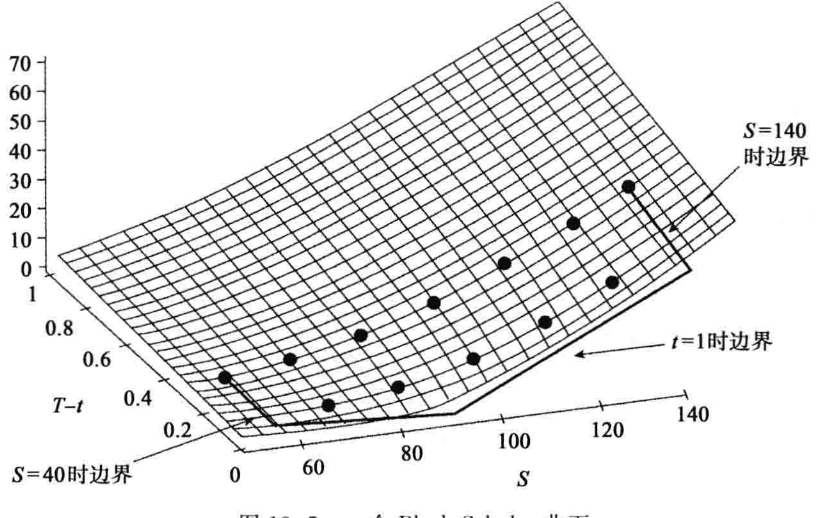


图 13-5　一个 Black-Scholes 曲面

这个差分问题可以采用递归的方法来求解，而递归的起点就是原偏微分方程的边界条件. 下一节我们将具体介绍如何利用边界条件.

13.7.3　边界条件

在上一节我们所介绍的差分方程中，一些 F_{ij} 的值是已知的. 例如，我们总可以知道在到期日期权价值关于 S_t 的函数. 对于 S_t 的极端值，也有一些在极限处成立的近似方法. 特别地：

- 对于非常高的 S_t，我们可以令 $S_t=S_{\max}$ 且
$$F(S_{\max}, t) \approx S_{\max} - K\mathrm{e}^{-r(T-t)} \tag{13.52}$$
这里，选择 $S_{\max}$ 使得赎回溢价与到期日的收益非常接近.
- 对于非常低的 S_t，我们令 $S_t=S_{\min}$ 且
$$F(S_{\min}, t) \approx 0 \tag{13.53}$$
这里，$S_{\min}$ 代表一个极低的价格. 看涨期权几乎没有可能在到期日变为实值. 最终的赎回溢价极可能为 0.
- 对于 $t=T$，我们可以确定
$$F(S_T, T) = \max[S_T - K, 0] \tag{13.54}$$

这些方法给出了 F_{ij} 在边界上的值. 在图 13-5 中，我们已将边界用实线标出来. 在方

225

程(13.47)中利用这些值，我们就能求出余下未知的 F_{ij}.

13.7.4 偏积分-微分方程数值解的技巧

我们仅关注偏积分-微分方程的积分部分，并给出一些求其数值解方面的技巧．这里我们考虑

$$\int_{-\infty}^{\infty}(F(S_{t-}\mathrm{e}^{y},t)-F(S_{t-},t))k(y)\mathrm{d}y \tag{13.55}$$

我们将要求数值解，其中列维密度是

$$k(y)=\frac{\mathrm{e}^{-\lambda_p y}}{\upsilon y}1_{y>0}+\frac{\mathrm{e}^{-\lambda_n |y|}}{\upsilon |y|}1_{y<0} \tag{13.56}$$

不失一般性，我们仅考虑正数 y.

$$\int_{0}^{\infty}(F(S_{t-}\mathrm{e}^{y},t)-F(S_{t-},t))k(y)\mathrm{d}y \tag{13.57}$$

通过变量代换 $x=\ln S$ 并且令 $\widetilde{F}(x, t)=F(S, t)$，可得

$$\int_{0}^{\infty}(\widetilde{F}(x+y,t)-\widetilde{F}(x,t))k(y)\mathrm{d}y \tag{13.58}$$

在 0 附近的区域情况比较特殊，对此我们需要另加考虑．首先研究对于某个很小的 $\Delta x>0$ 以下区域的积分：

$$\int_{0}^{\Delta x}(\widetilde{F}(x+y,t)-\widetilde{F}(x,t))k(y)\mathrm{d}y \tag{13.59}$$

利用泰勒展开，我们可以将 $\widetilde{F}(x+y,t)$表示为：

226

$$\widetilde{F}(x+y,t)=\widetilde{F}(x,t)+y\frac{\partial\widetilde{F}}{\partial x}(x,t)+\frac{1}{2}y^2\frac{\partial^2\widetilde{F}}{\partial x^2}(\xi,t) \tag{13.60}$$

或者等价地

$$\widetilde{F}(x+y,t)-\widetilde{F}(x,t)=y\frac{\partial\widetilde{F}}{\partial x}(x,t)+O(y^2) \tag{13.61}$$

将其代入(13.59)可得

$$\int_{0}^{\Delta x}(\widetilde{F}(x+y,t)-\widetilde{F}(x,t))k(y)\mathrm{d}y=\int_{0}^{\Delta x}y\frac{\partial\widetilde{F}}{\partial x}(x,t)k(y)\mathrm{d}y \tag{13.62}$$

$$=\frac{\partial\widetilde{F}}{\partial x}(x,t)\int_{0}^{\Delta x}y\frac{\mathrm{e}^{-\lambda_p y}}{\upsilon y}\mathrm{d}y \tag{13.63}$$

$$=\frac{1}{\upsilon}\frac{\partial\widetilde{F}}{\partial x}(x,t)\int_{0}^{\Delta x}\mathrm{e}^{-\lambda_p y}\mathrm{d}y \tag{13.64}$$

$$=\frac{1}{\upsilon\lambda_p}\frac{\partial\widetilde{F}}{\partial x}(x,t)(1-\mathrm{e}^{-\lambda_p\Delta x}) \tag{13.65}$$

观察任意一个不含 0 的小区间．任取整数 k，对于 $y\in(k\Delta x,(k+1)\Delta x)$，我们考虑积分：

$$\int_{k\Delta x}^{(k+1)\Delta x}(\widetilde{F}(x+y,t)-\widetilde{F}(x,t))k(y)\mathrm{d}y \tag{13.66}$$

利用线性近似，我们可以基于$(k\Delta x,(k+1)\Delta x)$将 $\widetilde{F}(x+y,t)$写为

$$\widetilde{F}(x+y,t)=\widetilde{F}(xf+k\Delta x,t)+\frac{\widetilde{F}(x+(k+1)\Delta x,t)-\widetilde{F}(x+k\Delta x,t)}{\Delta x}\times(y-k\Delta x)+O((\Delta x)^2) \tag{13.67}$$

代入(13.66)可得

$$\begin{aligned}
&\int_{k\Delta x}^{(k+1)\Delta x}\left(\widetilde{F}(x+k\Delta x,t)+\frac{\widetilde{F}(x+(k+1)\Delta x,t)-\widetilde{F}(x+k\Delta x,t)}{\Delta x}\times(y-k\Delta x)-\widetilde{F}(x,t)\right)\frac{e^{-\lambda_p y}}{\upsilon y}dy\\
&=(\widetilde{F}(x+k\Delta x,t)-\widetilde{F}(x,t)-k(\widetilde{F}(x+(k+1)\Delta x,t)-\widetilde{F}(x+k\Delta x,t)))\left(\int_{k\Delta x}^{(k+1)\Delta x}\frac{e^{-\lambda_p y}}{\upsilon y}dy\right)\\
&\quad+\frac{\widetilde{F}(x+(k+1)\Delta x,t)-\widetilde{F}(x+k\Delta x,t)}{\upsilon\Delta x}\left(\int_{k\Delta x}^{(k+1)\Delta x}e^{-\lambda_p y}dy\right)\\
&=\frac{(\widetilde{F}(x+k\Delta x,t)-\widetilde{F}(x,t)-k(\widetilde{F}(x+(k+1)\Delta x,t)-\widetilde{F}(x+k\Delta x,t)))}{\upsilon}\\
&\quad\times(\mathrm{ei}(k\Delta x\lambda_p)-\mathrm{ei}((k+1)\Delta x\lambda_p))+\frac{\widetilde{F}(x+(k+1)\Delta x,t)-\widetilde{F}(x+k\Delta x,t)}{\lambda_p\upsilon\Delta x}(e^{-\lambda_p k\Delta x}-e^{-\lambda_p(k+1)\Delta x})
\end{aligned} \tag{13.68}$$

其中

$$\mathrm{ei}(\xi)=\int_{\xi}^{\infty}\frac{e^{-t}}{t}dt \tag{13.69}$$

该积分是指数积分，可以通过相关程序快速计算．

13.8 小结

本章讨论了一些衍生资产定价中偏微分方程/偏积分-微分方程的例子．我们介绍了在看涨期权定价中引入第二个随机成分所带来的困难．我们还介绍了一些奇异期权以及偏微分方程/偏积分-微分方程相应的调整．

还有一个要点是 Black-Scholes 曲面的几何形状．我们可以从图 13-6 看到，标的资产价格的随机轨迹与曲面上衍生资产价格轨迹具有一一对应的关系．

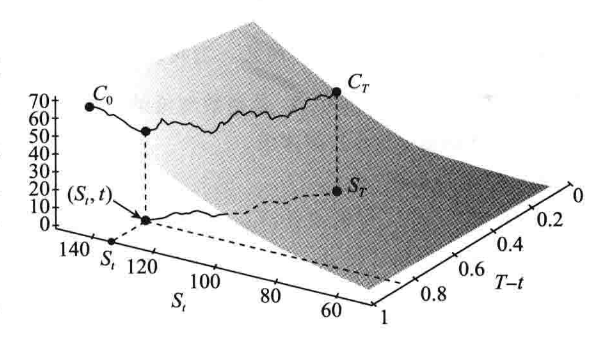


图 13-6 标的资产价格的随机轨迹所对应的 Black-Scholes 曲面上的随机路径

13.9 参考阅读

Ingersoll(1987)提供了一些利用偏微分方程进行资产定价的例子．在这一章中，我们的主要目的是通过一些例子来对偏微分方程作一个简单的介绍．如果读者需要更深入的了解，可以参考一些其他书籍．比如 Betounes(1998)是一本很好的参考读物，这本书以 MAPLE 为基础介绍了偏微分方程．

13.10 习题

227

这一节中的练习主要是为读者后 3 章的学习作铺垫，而非着重于计算偏微分方程．对偏微分方程有兴趣的读者可以从 Betounes(1998)找到一些有用的题目．

1. 设 X_t 为几何维纳过程，$X_t=e^{Y_t}$，其中 $Y_t\sim\mathcal{N}(\mu t,\sigma^2 t)$

根据定义

$$E[X_t\mid X_s,s<t]=E[e^{Y_t}\mid Y_s,s<t]$$

和下面这个显然成立的等式

$$E[e^{Y_t}\mid Y_s,s<t]=E[e^{(Y_t-Y_s)+Y_s}\mid Y_s,s<t]$$

来计算条件期望 $E[X_t\mid X_s,s<t]$

2. 本题主要练习如何得到鞅．假设 X_t 为几何过程，漂移参数为 μ，扩散参数为 σ.

(a) 什么情况下 $e^{-rt}X_t$ 是鞅？也就是说，什么情况下以下等式成立：

$$E[X_t\mid X_s,s<t]=e^{-rt}X_s$$

(b) 进一步，根据之前的推导

$$\begin{aligned}E[X_t\mid X_s,s<t]&=X_s e^{-rt}e^{(\mu+\frac{1}{2}\sigma^2)(t-s)}\\&=X_s e^{-rs}e^{-r(t-s)}e^{(\mu+\frac{1}{2}\sigma^2)(t-s)}\end{aligned}$$

也就是

$$E[X_t\mid X_s,s<t]=X_s e^{-rs}e^{(\mu+\frac{1}{2}\sigma^2)(t-s)}$$

是否可以选择适当的 μ，能够使 e^{-rtX_t} 成为鞅？

(c) 如果 $\mu=r+\sigma^2$，会怎么样？

(d) 现在令 $\mu=r-\frac{1}{2}\sigma^2$．注意，当我们选择不同定义的 μ 时，$e^{-rt}X_t$ 的分布也不同．

3. 考虑 $Z_t=e^{-rt}X_t$，其中 X_t 是指数型维纳过程：$X_t=e^{W_t}$.

(a) 计算增量 $dZ(t)$的期望值．

(b) Z_t 是否是鞅？

(c) 计算 $E[Z_t]$. 如何改变 X_t，才能使 Z_t 成为鞅？

(d) 在上一小题情况下，$E[Z_t]$会如何变化？

4. 考虑投资组合 $C(s,t)-\Delta S_t$. 假设 $S(t)$的动态变化满足：

$$dS_t=\mu S_t dt+\sigma dW_t \tag{13.70}$$

对 $C(s,t)-\Delta S_t$ 使用伊藤引理，可以得到这个投资组合的价值关于 dS 和 dt 的偏微分方程．令 Δ 为期权的 delta，即 $\Delta=\partial C/\partial S$. 请从一名投资者的角度，讨论这个投资组合价值关于时间的一阶偏导数(theta 值)与其关于标的资产价格的二阶偏导数(gamma 值)之间的关系．

228 ~ 229

5. 用 Matlab 写一个计算欧式看涨期权的定价函数．

第 14 章 衍生品定价：等价鞅测度

14.1 概率变换

衍生资产定价的最新方法未必用到无套利资产组合隐含的偏微分方程，而是通过使用 Girsanov 定理变换概率分布，将这类资产的价格转化为鞅进行计算.

这种方法与偏微分方程方法完全不同，它所使用的工具间接利用了无套利资产组合的存在，因此更难以直观地加以想象. 金融或经济专业的学生对这套新工具的熟悉程度可能弱于对偏微分方程方法的熟悉程度.

本章讨论的就是这些工具. 我们逐步展开，首先回顾一些简单概念和数学记号. 为了更好理解，我们还给出几个关于 Girsanov 定理应用的例子. 接下来给出该定理的完整表述. 随后，我们用一节的篇幅来说明该定理中用到的各种概念的直觉含义. 最后将该定理用到更复杂的例子中. 总的来说，本章的主要目的在于介绍变换概率分布这一概念，因而给出的金融市场实例会很少，相应的实例将会在下一章中讨论.

作为"测度"的概率

在 t 时刻，考虑服从正态分布且均值为 0、方差为 1 的随机变量 z_t，即

$$z_t \sim \mathcal{N}(0,1) \tag{14.1}$$

这个随机变量的概率密度函数 $f(z_t)$是

$$f(z_t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z_t^2}{2}} \tag{14.2}$$

假设我们想知道 z_t 落在某个特殊值 $\bar{z}$ 附近的概率. 那么先选取小区间 $\Delta>0$，然后计算正态密度函数在所求区域上的积分，即可得到该概率：

$$P\left(\bar{z}-\frac{1}{2}\Delta < z_t < \bar{z}+\frac{1}{2}\Delta\right) = \int_{\bar{z}-\frac{1}{2}\Delta}^{\bar{z}+\frac{1}{2}\Delta} \frac{1}{\sqrt{2\pi}} e^{-\frac{z_t^2}{2}} \mathrm{d}z_t \tag{14.3}$$

如果 $\bar{z}$ 附近的区域很小，$f(z_t)$的值在 z_t 从 $\bar{z}-\frac{1}{2}\Delta$ 变为 $\bar{z}+\frac{1}{2}\Delta$ 过程中不会发生太大变化. 也就是说我们可以在该区间内用 $f(\bar{z})$ 近似表示 $f(z_t)$，并将(14.3)右边的积分写成

$$\int_{\bar{z}-\frac{1}{2}\Delta}^{\bar{z}+\frac{1}{2}\Delta} \frac{1}{\sqrt{2\pi}} e^{-\frac{z_t^2}{2}} \mathrm{d}z_t \approx \frac{1}{\sqrt{2\pi}} e^{-\frac{\bar{z}^2}{2}} \int_{\bar{z}-\frac{1}{2}\Delta}^{\bar{z}+\frac{1}{2}\Delta} \mathrm{d}z_t \tag{14.4}$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{\bar{z}^2}{2}} \Delta \tag{14.5}$$

该结构如图 14-1 所示.(14.5)得到的概率等于底为 Δ、高为 $f(\bar{z})$ 的矩形(近似)面积.

从这个角度来看，概率相当于一个关于 z_t 在某个小区间内的可能取值的测度. 由于概率可以将任意集合映射到非负实数 $\mathbf{R}^{+}$ 上，因此概率是一种测度. 将 Δ 记作 dz_t，用 $dP(z_t)$ 或 dP(只要不会与基础随机变量发生混淆)表示测度：

232

$$dP(\bar{z}) = P\left(\bar{z} - \frac{1}{2}dz_t < z < \bar{z} + \frac{1}{2}dz_t\right) \tag{14.6}$$

可以认为该公式表示随机变量 z_t 落在中心为 $\bar{z}$、长度为 dz_t 的无穷小区间内的概率. 将不同 $\bar{z}$ 对应的概率相加得到的和可以用积分形式来表示：

$$\int_{-\infty}^{\infty} dP(z_t) = 1 \tag{14.7}$$

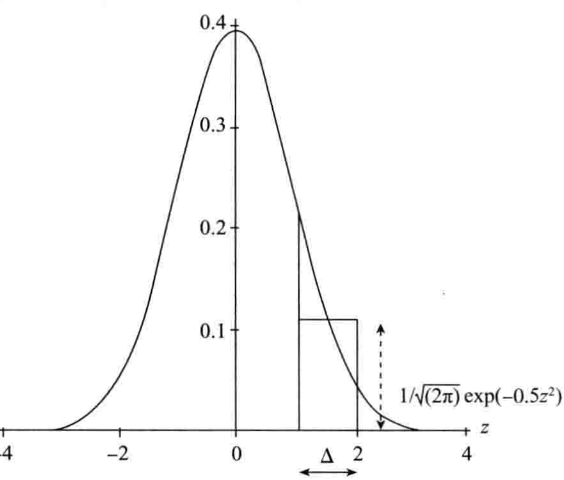


图 14-1 用单个矩形计算小区间上正态密度函数的积分

可以用类似的方法计算 z_t 的期望值

$$E[z_t] = \int_{-\infty}^{\infty} z_t dP(z_t) \tag{14.8}$$

可以将它看作 z_t 的平均值. 从几何图形上看，它确定了该概率值的中心点. 方差是另一种加权平均：

$$\begin{aligned} &E[z_t - E[z_t]]^2 \\ &= \int_{-\infty}^{\infty} [z_t - E[z_t]]^2 dP(z_t) \end{aligned} \tag{14.9}$$

方差也有其几何含义，它表示概率值如何在中心附近散布.

因此，当我们在某个特定概率测度 dP 下进行讨论时，通常已经知道随机变量密度函数的形状和位置.⊖

在这些条件下，我们可以对概率分布进行两种变换：

- 不改变分布形状，移动密度函数的位置. 图 14-2 将中心为

  $$\mu = -5 \tag{14.10}$$

  的正态密度函数变为中心为 0：

  $$\mu = 0 \tag{14.11}$$

  的正态密度函数.

- 也可以改变分布的形状. 一种方法是改变该分布的方差. 改变原始随机变量的大小可以做到这一点. 图 14-3 中随机变量 z_t 的方差从 4 减小到 1.

⊖ 本书中假设密度函数总是存在的. 在其他环境中，随机变量的密度函数可能不存在.

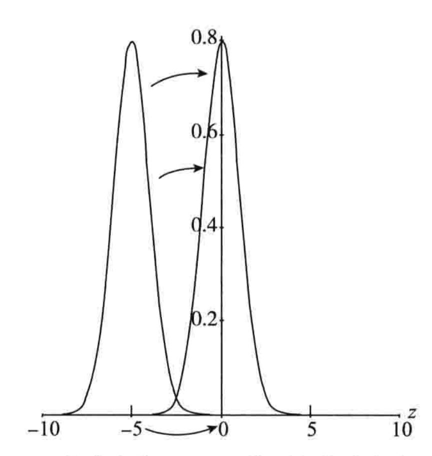


图 14-2　将中心为 $\mu=-5$ 的正态密度变成中心为 $\mu=0$ 的正态密度

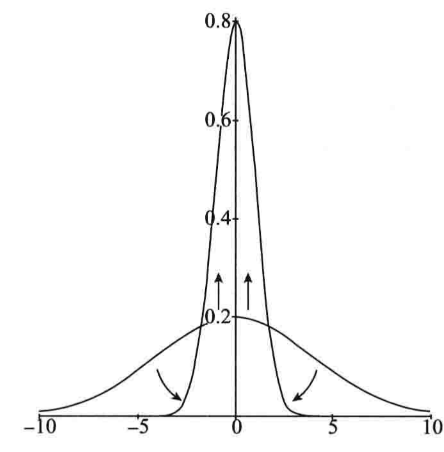


图 14-3　改变正态分布形状，随机变量方差从 4 变为 1

现在衍生资产定价的方法使用了一种新方法来变换概率测度 dP，改变随机过程 z_t 的均值．这种变换使具有正风险溢价的资产可以被视作为无风险资产来进行计算．本章讨论的就是这种方法．

我们将在下一节中讨论改变随机变量均值的两种不同方法．

14.2　改变均值

固定 t，则 z_t 是一个单变量随机变量．这时，我们有两种方法可以改变 z_t 的均值．第一种方法是对 z_t 本身进行变换．第二种方法与直觉相反，我们并不对 z_t 进行变换，而是对 z_t 所对应的概率测度进行变换．

这两种变换方式都可以在保留原始随机变量其他特征的同时，只改变其均值．然而在资产定价问题中，通常第一种方法不可行，所以第二种方法就变成了非常实用的工具．

接下来我们在单随机变量(而不是随机过程)环境下，讨论这两种方法变化的细节．对于更复杂的连续时间过程情况，将在 Girsanov 定理那一节中讨论．

14.2.1　方法 1：对变量本身进行变换

第一种可以改变随机变量均值的方法是标准变换，这种方法经常用于计量经济学和统计学中．将 z_t 加上常数 μ，得到新随机变量 $\tilde{z}_t=z_t+\mu$．[⊖]　233

这样定义的 $\tilde{z}_t$ 具有新的均值．

例如，如果原来

$$E[z_t]=0 \tag{14.12}$$

那么新随机变量有

⊖ μ 可以为负值．

$$E[\widetilde{z}_t] = E[z_t] + \mu = \mu \tag{14.13}$$

例 1

虽然这种变换非常简单，我们还是先来看一个例子，这对于后面的讨论也很重要．

假设随机变量 Z 按如下方式定义．掷骰子，按下面的规则定义 Z 的值

234

$$Z = \begin{cases} 10 & \text{掷出 1 或 2} \\ -3 & \text{掷出 3 或 4} \\ -1 & \text{掷出 5 或 6} \end{cases} \tag{14.14}$$

假设得到某个特定数字的概率为 1/6，通过计算 Z 的加权平均值，很容易得到出其均值：

$$E[Z] = \frac{1}{3}[10] + \frac{1}{3}[-3] + \frac{1}{3}[-1] \tag{14.15}$$

$$= 2 \tag{14.16}$$

假设我们想利用前面给出的方法改变 Z 的均值．更确切地说，是要得到方差相同但均值为 1 的新随机变量．称该新随机变量为 $\widetilde{Z}$，令

$$\widetilde{Z} = Z - 1 \tag{14.17}$$

利用均值公式，可以重新计算 $E[\widetilde{Z}]$

$$E[\widetilde{Z}] = \frac{1}{3}[10-1] + \frac{1}{3}[-3-1] + \frac{1}{3}[-1-1] \tag{14.18}$$

235

$$= 1 \tag{14.19}$$

从该变换中可以看出，为了改变 Z 的均值，我们对 Z 的可能取值进行变换，即将每个可能值减去 1，而其概率并未发生变化．

例 2

我们可以通过一个与金融关联性更强的例子来说明这种改变随机变量均值的方法．

固定 t，评级为 AAA 的公司债券的收益率 R_t 的期望值为

$$E[R_t] = r_t + E[\text{风险溢价}] \tag{14.20}$$

其中 r_t 表示期限相同的国债的无风险利率，这是一个已知的量．$E[\cdot]$表示在所有可能状态下的数学期望．用常数 α 表示风险溢价的期望值：

$$E[R_t] = r_t + \alpha \tag{14.21}$$

则 R_t 是一个均值为 $r_t + \alpha$ 的随机变量．

改变 R_t 均值的第一种方法是加上一个常数，从而得到新随机变量．该随机变量均值为

$$E[R_t + \mu] = r_t + \mu + \alpha \tag{14.22}$$

对于服从正态分布的随机变量，这种变换保持密度函数的形状不变，只移动分布中心的位置．图 14-2 给出了这种变换的图形．

如果 $\mu = -\alpha$，那么这种变换可以消掉 R_t 的风险溢价．注意，如果我们想用采用这种方法改变均值，必须先知道风险溢价 α．只有这样，我们才能从 R_t 中减去一个正确的量，

得到等价的无风险收益率.

本例很简单，没有给出进行这种均值变换的原因．下一个例子会对这一点做更多的说明.

例 3

我们先假设时间是离散的，用 $S_t(t=1, 2, \cdots)$来表示某种无红利金融资产的价格．显然在离散的时间点 t，$t+1$，…，S_t 是一个可以直接观察到的常量(而不是随机变量).

用 r_t 表示无风险收益率．风险资产 S_t 的收益率 R_t 需要大于 r_t，否则没有必要持有它.所以，在 t 时刻信息下的条件期望 $E[\cdot]$满足

$$E_t[S_{t+1}] > (1+r_t)S_t \tag{14.23}$$

即风险资产的期望增值速度通常要快于无风险资产．该式可以写作

$$\frac{1}{1+r_t}E_t[S_{t+1}] > S_t \tag{14.24}$$

左边表示将未来价格的期望值按无风险利率的贴现值．那么，我们可以认为存在 $\mu>0$ 使得

$$\frac{1}{1+r_t}E_t[S_{t+1}] = S_t(1+\mu) \tag{14.25}$$

可以将正常数 μ 或 $\mu+\mu r_t$ 解释为这项金融资产的风险溢价．我们也可以将(14.25)写成

$$\frac{E_t[S_{t+1}]}{S_t} = (1+r_t)(1+\mu) \tag{14.26}$$

等式左边的项 $E_t[S_{t+1}/S_t]$表示期望收益率 $E_t[1+R_t]$，即

$$E_t[1+R_t] = (1+r_t)(1+\mu) \tag{14.27}$$

这说明风险资产的预期收益率要比无风险收益率高约 μ:

$$E_t[R_t] \approx 1+r_t \tag{14.28}$$

这里 r_t 和 μ 足够小，因此它们的交叉乘积可以忽略.

236

在这样的条件下，μ 为持有资产一段时间的风险溢价，r_t 为无风险贴现因子.

金融分析师为了得到该资产当下的公平市场价格，需要计算 S_t. 其中一种计算方法是利用关系式

$$E_t\left[\frac{1}{1+R_t}S_{t+1}\right] = S_t \tag{14.29}$$

来计算左边的数学期望.[1]

但这种方法需要知道 R_t 的分布，而这又需要知道风险溢价 μ.[2]然而在得到公平市场价值 S_t 前就知道风险溢价的情况很罕见，因此使用(14.29)并不能让我们得到 S_t 的值.[3]

---

[1] 该关系式即收益率 R_t 的定义．将下一阶段的价格除以 $1+R_t$，自然可以得到现在的价值.

[2] 只有已知 μ 才能计算 R_t 的均值，R_t 的分布才能随之固定.

[3] 除此以外还有一个难题.(14.29)左边项是 R_t 的非线性函数，因此我们不能直接将期望运算符移到 R_t 前:

$$E\left[\frac{1}{1+R_t}S_{t+1}\right] \neq \left[\frac{1}{1+E[R]_t}S_{t+1}\right] \tag{14.30}$$

这使计算更加复杂.

另一方面，如果可以在不使用 μ 的前提下变换 R_t 的均值，这种方法可能有用.[⊖] 为此我们必须找到另一种变换 R_t 分布的方法.

假如在另一种概率分布 Q 下，我们可以得到新期望值

$$E_t^Q\left[\frac{1}{1+r_t}S_{t+1}\right]=S_t \tag{14.31}$$

那么这种方法就可以用于计算 S_t. 当使用该公式计算 S_t 时，我们可以先用描述 S_t 价格过程的模型预测出 S_{t+1}，然后用(已知的) r_t 对平均预测值进行贴现，就得到 S_t 的预测值.

$E_t^Q[\cdot]$和 r_t 在这里有什么含义呢？r_t 仍表示无风险利率. 该期望运算是建立在风险中性概率条件下的. 通过这些变换，我们可以消去 R_t 中的风险溢价：

$$R_t-\mu=r_t \tag{14.32}$$

这个变换的巧妙之处在于我们可以在不用到 μ 的情况下完成对均值的变换. 虽然乍看起来这是一项不可能完成的任务，但这第二种变换方法确实做到了这一点.

14.2.2 方法 2：对概率进行运算

改变随机变量均值的第二种变换方法就是保持随机变量本身不变，只对 z_t 对应的概率测度进行变换. 我们通过一系列的例子来介绍这种方法. 最后我们给出 Girsanov 定理，将这种测度变换的方法推广到连续时间的随机过程. 这种方法看上去可能是不那么直观，但实际上却非常简单，我们会在例 1 中看到这一点.

例

我们仍然考虑前一节中第一个例子，将 Z 定义为掷骰子结果的函数：

$$Z=\begin{cases}10 & \text{掷出 1 或 2}\\ -3 & \text{掷出 3 或 4}\\ -1 & \text{掷出 5 或 6}\end{cases} \tag{14.33}$$

前面已经计算出它的均值

237

$$E[Z]=2 \tag{14.34}$$

及方差

$$V[Z]=E[Z-E[Z]]^2=\frac{1}{3}[10-2]^2+\frac{1}{3}[-3-2]^2+\frac{1}{3}[-1-2]^2=\frac{98}{3} \tag{14.35}$$

假设我们要对该随机变量进行变换，使其方差不变，均值变为 1.

我们现在对掷骰子结果原始概率进行如下变换：

$$P(\text{得到 1 或 2})=\frac{1}{3}\rightarrow Q(\text{得到 1 或 2})=\frac{122}{429} \tag{14.36}$$

$$P(\text{得到 3 或 4})=\frac{1}{3}\rightarrow Q(\text{得到 3 或 4})=\frac{22}{39} \tag{14.37}$$

⊖ 因为 R_t 分布的均值可能等于 r_t.

$$P(\text{得到 5 或 6}) = \frac{1}{3} \rightarrow Q(\text{得到 5 或 6}) = \frac{5}{33} \tag{14.38}$$

这里，我们用 Q 来标记这个新得到的概率测度．

现在重新计算新概率下的均值：

$$E^Q[Z] = \left[\frac{122}{429}\right][10] + \frac{22}{39}[-3] + \frac{5}{33}[-1] = 1 \tag{14.39}$$

均值确实为 1. 计算其方差：

$$E^Q[Z]^2 = \left[\frac{122}{429}\right][10-1]^2 + \frac{22}{39}[-3-1]^2 + \frac{5}{33}[-1-1]^2 = \frac{98}{3} \tag{14.40}$$

方差没有改变．(14.36)－(14.38)中的概率变换完全达到了第一种变换方法所达到的效果，但这种方法是对概率测度 $P(Z)$进行变换而不是 Z 本身进行变换．

这里我们需要要强调一点，这个新得到的概率与该试验的真实概率无关．掷骰子的真实概率仍用原始概率 P 表示．

读者可能已经注意到我们所使用的记法．实际上，我们要将新的期望运算符写做 $E^Q[\cdot]$而不是 $E[\cdot]$. 这是因为计算均值时使用的概率不再是 P，如果仍然使用 $E[\cdot]$会造成混淆．因此，使用这种方法时，计算期望值要特别注意标记所使用的概率．[⊖]

14.3 Girsanov 定理

显然，刚刚讨论的例子都是经过简化的：第一，我们讨论的随机变量其状态空间有限——随机变量的可能取值范围是有限个元素组成的集合；第二，我们讨论的是单个随机变量，而不是一个随时间变化的随机过程．

Girsanov 定理为等价测度变换给出了一个通用方法．它适用于布朗运动，因此其对应的状态空间是连续的，进一步地说，这个变换可以被推广到连续时间随机过程中．

这里，我们称变换前后的概率测度是“等价”的，是因为如果某集合在其中一个概率测度下的测度值不为 0，那么在它在另一个概率测度下的测度值也不为 0(关于等价测度，我们将在后面有更详细的讨论). 因此，尽管两种概率分布不同，我们仍然可以通过适当变换将一种测度变为另一种测度．同时，利用这种对应性，我们可以在计算中使用某种简单的分布，然后，如果需要，我们还可以将它变回原分布．

238

相应地，如果我们计算某个数学期望时，发现在某个等价测度下进行计算更简单，虽然这个新测度不是真实测度，我们仍然应该先做概率测度变换．这是因为，我们的目的不是描述各种状态的(真实)概率值，而只是希望以一种更便捷的方式来进行计算期望值．

这种通用方法可以总结如下：(1)需要计算某个数学期望．(2)变换原始概率测度，使

⊖ 一些读者可能想知道如何得出新概率．在这种特殊情况下很简单．将概率看做未知，通过三个条件求出概率．第一个条件是概率和为 1. 第二个条件是新均值为 1. 第三个条件是方差为$\frac{98}{3}$.

期望计算更容易.(3)在新概率下计算数学期望.(4)得出结果后，如果需要，再将该概率变回原始分布.

现在，我们在更复杂的环境中讨论这种概率变换.通过一系列越来越复杂的例子来引入 Girsanov 定理，最后给出这条定理的一般形式，并讨论它的假设条件和含义.

14.3.1 正态分布的随机变量

固定 t，考虑正态随机变量 z_t：

$$z_t \sim \mathcal{N}(0,1) \tag{14.41}$$

写出 z_t 的密度函数 $f(z_t)$在概率测度 P 下的形式，

$$\mathrm{d}P(z_t) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(z_t)^2}\mathrm{d}z_t \tag{14.42}$$

尽管我们在这个例子中讨论的仍然是单个随机变量，而不是讨论一个随机过程，但这里这个随机变量的状态空间仍是连续的，因此情况比我们之前研究过的掷骰子的例子要更复杂.

接下来，我们定义函数

$$\xi(z_t) = \mathrm{e}^{z_t\mu-\frac{1}{2}\mu^2} \tag{14.43}$$

将 $\xi(z_t)$与 $\mathrm{d}P(z_t)$相乘，我们可以得到一个新概率测度：

$$[\mathrm{d}P(z_t)][\xi(z_t)] = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(z_t)^2+\mu z_t-\frac{1}{2}(\mu)^2}\mathrm{d}z_t \tag{14.44}$$

将指数中的项进行化简，得到

$$\mathrm{d}Q(z_t) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(z_t-\mu)^2}\mathrm{d}z_t \tag{14.45}$$

显然 $\mathrm{d}Q(z_t)$是由

$$\mathrm{d}Q(z_t) = \mathrm{d}P(z_t)\xi(z_t) \tag{14.46}$$

定义的新概率测度.

从(14.45)的密度函数中可以看出，z_t 在这个新的概率测度下仍然是一个正态分布的随机变量，且它的均值为 μ，方差为 1.

这里，我们通过将 $\xi(z_t)$与 $\mathrm{d}P(z_t)$相乘变换到 Q 概率测度，成功地改变了 z_t 的均值.在这个特殊例子中，我们注意到乘以 $\xi(z_t)$不改变原概率测度的形状(即并没有改变其方差). 事实上，(14.45)的图形仍是一个方差为 1 的钟形高斯曲线.但 $P(z_t)$与 $Q(z_t)$是两种不同的测度.两种测度下均值不同，它们对 z 轴上区间所赋权重不同.

在 $P(z_t)$测度下，随机变量 z_t 均值为 0，$E^P[z_t]=0$,，方差 $E^P[z_t^2]=1$. 然而在 $Q(z_t)$测度下，z_t 均值为 $E^Q[z_t]=\mu$，方差仍为 1.

本例说明存在 $\xi(z_t)$，将它与原始概率相乘可以得到新概率.随机变量在这个新测度下仍服从正态分布，只是均值不同.

最后，这个改变均值的测度变换

$$dQ(z_t) = \xi(z_t)dP(z_t) \tag{14.47}$$

是可逆的：

$$\xi(z_t)^{-1}dQ(z_t) = dP(z_t) \tag{14.48}$$

同时，这个变换不会改变 z_t 的方差，并且在已知 μ 和 σ 时是唯一的．

现在我们可以总结一下两种改变均值的变换方法：[⊖]

- 方法 1：减去均值．已知随机变量：

$$Z \sim \mathcal{N}(\mu,1) \tag{14.49}$$

将新随机变量 $\widetilde{Z}$ 定义为 Z 的变换：

$$\widetilde{Z} = \frac{Z-\mu}{1} \sim \mathcal{N}(0,1) \tag{14.50}$$

则 $\widetilde{Z}$ 均值为 0.

- 方法 2：使用等价测度．已知 P 测度下随机变量：

$$Z \sim P = \mathcal{N}(\mu,1) \tag{14.51}$$

将 $\xi(Z)$ 乘以 $d(P)$，将 $d(P)$ 变为新测度 $d(Q)$，使得：

$$Z \sim Q = \mathcal{N}(0,1) \tag{14.52}$$

我们接下来要讨论的是，对于一列正态随机变量序列 z_1，z_2，…，z_t，是否可以得到相同的变换．

14.3.2 正态随机向量

前面的例子说明了如何利用函数 $\xi(z_t)$ 乘以相应的概率测度，改变正态随机变量的均值．变换后的测度是另一种概率分布，它使 z_t 的均值发生变化，而方差不变．

如果已知的是服从多元正态分布的随机向量，我们还可以通过类似方法处理吗？

答案是肯定的．为了方便起见，我们仅对二维情况进行证明，推广到 n 维方法类似．

固定 t，假设已知随机变量 z_{1t}、z_{2t} 服从二元正态分布．其密度函数为：

$$f(z_{1t},z_{2t}) = \frac{1}{2\pi\sqrt{|\Omega|}} e^{-\frac{1}{2}[(z_{1t}-\mu),(z_{2t}-\mu)]'\Omega^{-1}[(z_{1t}-\mu),(z_{2t}-\mu)]} \tag{14.53}$$

其中 Ω 是随机向量 $[(z_{1t}-\mu),(z_{2t}-\mu)]$ 的方差-协方差矩阵

$$\Omega = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix} \tag{14.54}$$

其中 $\sigma_i^2(i=1,2)$ 表示方差，σ_{12} 表示 z_{1t}，z_{2t} 的协方差，$|\cdot|$ 表示行列式：

$$|\Omega| = \sigma_1^2\sigma_2^2 - \sigma_{12}^2 \tag{14.55}$$

最后 μ_1，μ_2 是 z_{1t}，z_{2t} 的均值．

⊖ 我们对记号稍作简化．

这个二维随机向量的联合概率测度可以用下式表示

$$dP(z_{1t},z_{2t}) = f(z_{1t},z_{2t})dz_{1t}dz_{2t} \tag{14.56}$$

这个表达式是用小矩形面积表示的概率，其中矩形中心特殊值为 z_{1t}、z_{2t}. 它给出了 z_{1t}、z_{2t}落在这个特定矩形内的概率.

假设我们想要将 z_{1t}，z_{2t}的均值从 μ_1，μ_2 变为 0，同时不改变方差．可以像前面的例子一样，通过乘以函数 $\xi(z_{1t},z_{2t})$改变概率 $dP(z_{1t},z_{2t})$，做到这一点吗？

240

答案是肯定的．考虑函数

$$\xi(z_{1t},z_{2t}) = e^{-\frac{1}{2}[z_{1t},z_{2t}]'\Omega^{-1}[\mu_1,\mu_1]+\frac{1}{2}[\mu_1,\mu_1]'\Omega^{-1}[\mu_1,\mu_1]} \tag{14.57}$$

我们可以利用它定义一个新概率测度 $Q(z_{1t},z_{2t})$

$$dQ(z_{1t},z_{2t}) = \xi(z_{1t},z_{2t})dP(z_{1t},z_{2t}) \tag{14.58}$$

可以用(14.57)中的 $\xi(z_{1t},z_{2t})$乘以(14.53)得到 $dQ(z_{1t},z_{2t})$. 这两项的积为

$$dQ(z_{1t},z_{2t}) = \frac{1}{2\pi\sqrt{|\Omega|}}e^{-\frac{1}{2}[z_{1t},z_{2t}]'\Omega^{-1}[z_{1t},z_{2t}]}dz_{1t}dz_{2t} \tag{14.59}$$

可以看出这就是均值为 0、方差一协方差矩阵为 Ω 的二维正态分布随机向量的密度函数. 乘以 $\xi(z_{1t},z_{2t})$达到了想要的效果．通过对原概率进行变换，将二维向量的非零均值变为 0.

本例讨论的是二维随机向量．如果已知的是 k 维正态随机向量序列$[z_{1t},z_{2t},\cdots,z_{kt}]$，仍然可以使用相同变换．只需要改变(14.53)及(14.57)中相应向量和矩阵的维度.

注释

在继续讨论前，我们想要强调一点规律性，读者们可能已经观察到了.

不管 z_t 是长度为 k 的随机向量，还是一元随机变量，在将概率测度 $P(z_t)$变换成 $Q(z_t)$时，我们都用到了函数 $\xi(z_t)$. 该函数具有下面的结构

$$\xi(z_t) = e^{-\frac{1}{2}z_t'\Omega^{-1}\mu+\frac{1}{2}\mu'\Omega^{-1}\mu} \tag{14.60}$$

在标量情况下(即一维的情况)变为

$$\xi(z_t) = e^{-\frac{1}{2}\frac{z_t\mu}{\sigma^2}+\frac{1}{2}\frac{\mu^2}{\sigma^2}} \tag{14.61}$$

现在我们讨论该函数形式是如何产生的．在正态分布中，参数 μ 表示均值，仅出现在 e 的指数中，而且该指数具有平方形式：

$$-\frac{1}{2}\frac{(z_t-\mu)^2}{\sigma^2} \tag{14.62}$$

为了将该表达式变为

$$-\frac{1}{2}\frac{z_t^2}{\sigma^2} \tag{14.63}$$

我们只需要加上

$$-\frac{z_t\mu+1/2\mu^2}{\sigma^2} \tag{14.64}$$

这就决定了函数 $\xi(z_t)$的形式．将原始概率测度乘以 $\xi(z_t)$便可以使 e 的指数完成上面的变换.

到此为止，读者可能想知道我们是否会对 $\xi(z_t)$赋予更深的含义．下一节讨论的就是这一点.

14.3.3 Radon-Nikodym 导数

再次考虑函数 $\xi(z_t)$，其中 $\sigma=1$：[⊖]

$$\xi(z_t) = e^{-\mu z_t+\frac{1}{2}\mu^2} \tag{14.65}$$

241

利用 $\xi(z_t)$将 $dP(z_t)$变为新概率测度 $dQ(z_t)$：

$$dQ(z_t) = \xi(z_t)dP(z_t) \tag{14.66}$$

或者两边同时除以 $dP(z_t)$

$$\frac{dQ(z_t)}{dP(z_t)} = \xi(z_t) \tag{14.67}$$

可以将该表达式看成为一个导数的表达式，表示新测度关于 P 测度的导数，记为 $\xi(z_t)$. 这种导数叫做 Radon-Nikodym 导数，我们也可以认为 $\xi(z_t)$是 Q 测度关于 P 测度的密度函数．

据此，如果关于 P 概率的 Radon-Nikodym 导数存在，我们就可以用得到的 $\xi(z_t)$变换 z_t 的均值，保持方差结构不变．

这种变换对于金融市场参与者非常有意义，因为它可以消掉资产价格的风险溢价，而保持波动率结构不变．对于期权来说，其价值不依赖于标的资产价格增长的均值，而是主要由标的资产的波动率决定．在这种情况下，利用$\xi(z_t)$变换原始概率分布将非常方便.

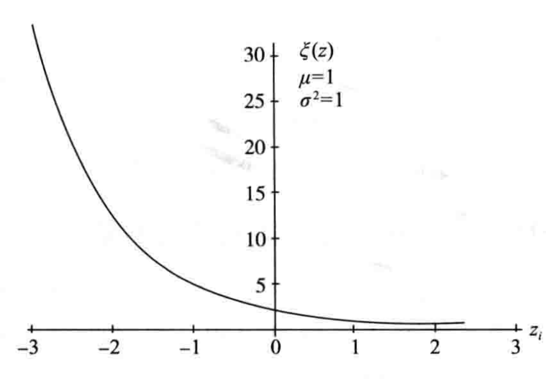


图 14-4 Radon-Nikodym 导数图像

在图 14-4 中，我们给出了 $\xi(z_t)$的一个例子.

⊖ 函数

$$\xi(z_t)=e^{-\mu z_t+\frac{1}{2}\mu^2} \tag{14.68}$$

将 z_t 的均值变为 0，而函数

$$\xi(z_t)^{-1}=e^{\mu z_t-\frac{1}{2}\mu^2} \tag{14.69}$$

将 Z 的原始均值 0 变为 μ.

14.3.4 等价测度

Radon-Nikodym 导数

$$\frac{dQ(z_t)}{dP(z_t)} = \xi(z_t) \tag{14.70}$$

在什么时候存在呢？即我们什么时候可以进行下面的变换呢？

$$dQ(z_t) = \xi(z_t)dP(z_t) \tag{14.71}$$

为了使比率

$$\frac{dQ(z_t)}{dP(z_t)} \tag{14.72}$$

有意义，分母上的概率不能为 0. 为了进行反向运算，分子也不能为 0. 而分子分母都是无穷小区间 dz 对应的概率．因此，要使 Radon-Nikodym 导数存在，当 dz 在 Q 测度下概率非零时，它在 P 测度下概率也必须非零，反之亦然．也就是说：

条件：已知区间 dz_t，概率测度 P、Q 满足

$$Q(dz_t) > 0 \quad 当且仅当 \quad P(dz_t) > 0 \tag{14.73}$$

若满足该条件，则存在 $\xi(z_t)$，进而我们就可以在两种测度之间进行相互变换，关系式为

$$dQ(z_t) = \xi(z_t)dP(z_t) \tag{14.74}$$

及

$$dP(z_t) = \xi(z_t)^{-1}dQ(z_t) \tag{14.75}$$

这意味着对于任意实际用途，这两种测度都是等价的．

242

因此它们被称为等价概率测度．

14.4 Girsanov 定理的内容

至此，给出的例子在连续时间金融中的应用都具有一些限制性．连续时间金融讨论的是连续或右连续随机过程，而到此为止的变换都只是对有限随机变量序列的变换. Girsanov 定理给出了当 z_t 为连续随机过程时，Radon-Nikodym 导数 $\xi(z_t)$存在的条件．在连续时间金融中，我们就是用这条定理来进行随机变换．

我们首先给出正式的 Girsanov 定理，随后对其进行延伸讨论．

Girsanov 定理的条件如下：已知区间$[0,T]$上的一簇信息集$\{I_t\}$，T 有限．[㊀]

在这段区间上定义随机过程 ξ_t：

$$\xi_t = e^{\left(\int_0^t X_u dW_u - \frac{1}{2}\int_0^t X_u^2 du\right)}, t \in [0,\infty) \tag{14.76}$$

其中 X_t 是 I_t- 可测过程[㊁]. W_t 是概率测度 P 下的维纳过程．

㊀ 对于金融衍生品来说，这不是一个非常严格的限制．几乎所有金融衍生品到期日都是有限的．通常来说，衍生品的期限都少于 1 年．

㊁ 此即，给定信息集 I_t 就能准确知道 X_t 的值．

我们对 X_t 加一个条件，使 X_t 变化不要过于剧烈：

$$E[e^{\int_0^t X_u^2 du}] < \infty, t \in [0, \infty) \tag{14.77}$$

也就是说 X_t 不会随时间产生太过剧烈的增加(或减少). 我们称(14.77)为 Novikov 条件.

在连续时间条件下，密度 ξ_t 的一条新性质变得非常重要：只要满足 Novikov 条件，ξ_t 就是一个平方可积鞅．我们首先证明这一点．

利用伊藤引理，我们可以得到

$$d\xi_t = \left[e^{\left(\int_0^t X_u dW_u - \frac{1}{2}\int_0^t X_u^2 du\right)} \right] [X_t dW_t] \tag{14.78}$$

这个微分可以化简为

$$d\xi_t = \xi_t X_t dW_t \tag{14.79}$$

令(14.76)中 $t=0$，有

$$\xi_0 = 1 \tag{14.80}$$

对(14.79)进行随机积分，得到

$$\xi_t = 1 + \int_0^t \xi_s X_s dW_s \tag{14.81}$$

但

$$\int_0^t \xi_s X_s dW_s \tag{14.82}$$

是关于维纳过程的随机积分．而且 $\xi_s X_x$ 项是 I_t 适应的，不会迅速变化．和第 6 章中的讨论一样，这些都表明这个积分是一个(平方可积)鞅

$$E\left[\int_0^t \xi_s X_s dW_s \mid I_u\right] = \int_0^t \xi_s X_s dW_s \tag{14.83}$$

其中 $u < t$. (14.81)说明 ξ_t 是(平方可积)鞅．

现在我们可以正式给出 Girsanov 定理了．

定理 6 如果随机过程 ξ_t 定义为(14.76)，在概率测度 P 下是关于信息集 I_t 可测的鞅，定义 W_t^*

$$W_t^* = W_t - \int_0^t X_u dW_u, \quad t \in [0, \infty) \tag{14.84}$$

那么 W_t^*在概率测度 Q 下是关于 I_t 可测的维纳过程，其中 Q 测度为

$$Q(A) = E^P(1_A \xi_T) \tag{14.85}$$

其中 A 为由 I_T 决定的事件，1_A 为该事件的指示函数．

243

这个定理说明，只要已知维纳过程 W_t，将该过程的概率分布乘以 ξ_t，就可以得到 Q 测度下的新维纳过程．这两个过程通过下式相互关联

$$dW_t^* = dW_t - X_t dt \tag{14.86}$$

此即，从 W_t 中减去 I_t 适应的漂移项就可以得到该等式．

做这样的变换的核心条件在于 ξ_t 是鞅，且 $\xi_0 = 1$.

现在我们对 Girsanov 定理的记法和假设进行详细讨论．定理的证明见 Liptser 和

Shiryayev(1977).

14.5 关于 Girsanov 定理的讨论

我们在本节对 Girsanov 定理用到的记法和假设条件进行系统讨论，并将它们与前面的例子联系起来．我们还将说明它们与金融模型中概念的相关性．

我们从函数 ξ_t 开始讨论：

$$\xi_t = e^{\frac{1}{\sigma^2}\left[\int_0^t X_u dW_u - \frac{1}{2}\int_0^t X_u^2 du\right]} \tag{14.87}$$

这里我们提出积分中的(常数)σ^2 项作为因子．该项也可以包含在 X_u 中．

假设积分中的 X_u 恒等于常数 μ：

$$X_u = \mu \tag{14.88}$$

化简指数中的积分，由于 $W_0=0$，我们可以得到

$$\xi_t = e^{\frac{1}{\sigma^2}\left[\mu W_t - \frac{1}{2}\mu^2 t\right]} \tag{14.89}$$

该等式与前面讨论的 ξ_t 类似．这说明了以下非常重要的几点：

1) Girsanov 定理中用到的符号 X_t 与简单环境中 μ 所起到的作用是相同的．它度量了原始均值变化的幅度．

2) 在前面的例子中，μ 是独立于时间的．而这里 X_t 可以依赖于任意随机量，只要该随机量在 t 时刻时已知．这就是令 X_t 关于 I_t 适应的含义．因此可以对漂移项进行更复杂的变换．

3) ξ_t 是鞅，且 $\xi_0=1$.

接下来考虑维纳过程．关于该过程，有几点不那么直观．W_t 和 W_t^* 都是标准维纳过程，因此均不含任何漂移项．但它们通过下式相互关联

$$dW_t^* = dW_t - X_t dt \tag{14.90}$$

也就是说只要 X_t 不等于零，这两个过程中就至少有一个具有非零漂移项．如何解释该矛盾呢?

重点在于 W_t^* 在 Q 测度下具有零漂移项，而 W_t 在 P 测度下具有零漂移项．因此如果我们将概率测度从 P 变换到 Q，可以用 W_t^* 表示在这个新概率测度下，动态系统中不可预测的误差．

同样地，由于包含项 $-X_t dt$，用 W_t^* 替代 W_t 做误差项可以准确地将原始 SDE 中漂移项减少 $-X_t dt$. 因此，如果将 X_t 看做依赖于时间的风险溢价，这个变换可以使所有风险资产都按无风险利率增长．

最后，考虑关系式

$$Q(A) = E^P(1_A \xi_T) \tag{14.91}$$

1_A 的含义是什么？我们如何推出该关系式?

示性函数 1_A 为当 A 事件发生时取值为 1 的函数．我们可以将前面的等式重新写成

244

$$Q(A) = E^P(1_A \xi_T) = \int_A \xi_T dP \tag{14.92}$$

当 A 为无穷小区间时，我们就有

$$dQ = \xi_T dP \tag{14.93}$$

这类似于前面简单环境中的概率变换．

在 SDE 中的应用

我们给出一个具有启发性的例子．

令 dS_t 表示股票价格发生的无穷小变动．假设这些变动由服从正态分布的无穷小振动所驱动，那么我们可以用关于维纳过程的随机微分方程表示 S_t

$$dS_t = \mu dt + \sigma dW_t, t \in [0, \infty) \tag{14.94}$$

其中 $W_0 = 0$.[⊖]假设 W_t 对应概率分布为 P，且

$$dP(W_t) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2t}(W_t)^2} dW_t \tag{14.95}$$

如果漂移项 μdt 不为零，我们接下来看 S_t 是不是一个鞅．首先我们知道

$$S_t = \mu\int_0^t ds + \sigma\int_0^t dW_s, t \in [0, \infty) \tag{14.96}$$

其中 $S_0 = 0$. 这等价于

$$S_t = \mu t + \sigma W_t \tag{14.97}$$

因此

$$E[S_{t+s} | S_t] = \mu(t+s) + \sigma E[W_{t+s} - W_t | S_t] + \sigma W_t \tag{14.98}$$

$$= S_t + \mu s \tag{14.99}$$

由于当给定 S_t 时，$W_{t+s} - W_t$ 是不可预测的．所以，对任意 $\mu > 0$，$s > 0$：

$$E[S_{t+s} | S_t] > S_t \tag{14.100}$$

这个严格不等式说明 S_t 不是鞅．

但我们可以通过消去其漂移项，来将 S_t 变为鞅．

其中一种消除漂移项方法在前面讨论过——从 S_t 中减去适当的均值，并定义

$$S_t^* = S_t - \mu t \tag{14.101}$$

则S_t^* 为鞅．

这种变换的缺点在于，为了得到 S_t^*，需要先知道 μ. 但 μ 包含了风险股票收益率具有的全部风险溢价．一般来说，这种风险溢价在得到资产公平市场价格前是未知的．

第二种将 S_t 变为鞅的方法更有效．利用 Girsanov 定理，我们可以轻松地变换到等价测度 Q，使 S_t 的漂移项变为 0.

为此，我们只需要找到函数 $\xi(S_t)$，将它乘以 S_t 的原始概率测度．这样，S_t 在 P 测度下可能是下鞅

⊖ 该等式再次使价格为负的可能性大于零．我们使用它的原因在于书写简便．我们大部分情况是用几何 SDE 来描述价格过程，这将在下一章中讨论．

$$E^P[S_{t+s} \mid S_t] > S_t \tag{14.102}$$

但它在 Q 测度下是鞅

$$E^Q[S_{t+s} \mid S_t] = S_t \tag{14.103}$$

同往常一样，运算符 $E[\cdot \mid \cdot]$的上标表示用于计算数学期望的概率测度．

为了进行该变换，我们需要先计算函数 $\xi(S_t)$．首先，S_t 的密度函数为

$$f_s = \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-\frac{1}{2t\sigma^2}[S_t - \mu t]^2} \tag{14.104}$$

它定义在 P 概率测度下．

我们要得到 Q 概率测度，使得 S_t 在 Q 测度下为鞅．定义

245

$$\xi(S_t) = e^{-\frac{1}{\sigma^2}[\mu t S_t - \frac{1}{2}\mu^2 t]} \tag{14.105}$$

f_s 乘以 $\xi(S_t)$得到

$$dQ(S_t) = \xi(S_t) dP(S_t) = e^{-\frac{1}{\sigma^2}[\mu t S_t - \frac{1}{2}\mu^2 t]} \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-\frac{1}{2t\sigma^2}[S_t - \mu t]^2} dS_t \tag{14.106}$$

整理指数项，我们得到

$$= \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-\frac{S_t^2}{2t\sigma^2}} dS_t \tag{14.107}$$

这是关于漂移项为 0、扩散项为 σ 的正态分布过程的概率测度．这意味着我们可以将 S_t 的增量写成新形式：

$$dS_t = \sigma dW_t^* \tag{14.108}$$

在这里，这样的 S_t 是鞅．维纳过程定义在 Q 概率测度下．

14.6 选择哪种概率

即使在这种水平的讨论中，合成概率 Q 仍对金融证券定价起着核心作用．根据第 2 章中的讨论，在无套利条件下，离散时间下任意一个在流动市场中的证券的公平价格都可以写成鞅等式：

$$S_t = E^Q[D_t S_T] \tag{14.109}$$

其中 $t<T$，D_t 为贴现因子，它可以是已知的或者随机的．如果不考虑外汇兑换或者股利支出，我们可以将储蓄账户正则化，此时 D_t 是无风险利率 r_t 的函数．如果 $r_t = r$ 为常数，那么 D_t 已知，可以提出期望运算符作为一个常数因子．

当 D_t 已知，概率测度 Q 已知时，等式(14.109)是非常有用的分析工具．因为所有衍生资产都存在到期时间 T，使得衍生资产价格 C_T 根据合约依赖于 S_T，那么当 $F(\cdot)$已知时，由$C_t = F(S_t, t)$有：

$$C_t = E^Q[D_t C_T] = E^Q[D_t F(S_T, T)]$$

从业者接下来需要进行下面简单的几步，来对衍生合约进行定价：

- 首先，需要选择概率分布 Q. 这项工作通常通过提取标的过程的一阶矩和二阶矩间接

完成，金融基本定理揭示了这一点．例如，对于不产生支付、不涉及外汇的证券，令小区间 $\Delta>0$ 有：

$$\frac{E^Q[S_{t+\Delta}]}{S_t}\approx r_t\Delta \tag{14.110}$$

它确立了所设随机微分方程的无套利条件．

- 其次，市场中的从业者需要校对 SDE 中波动率参数．这项工作并不简单，通常建立在存在流动期权或利率上限/利率下限市场的基础上，这些市场能够直接提供波动率数据．但即使这样，对波动率的校对也要非常谨慎．
- 只要确定了合适的概率测度变换和正确的价格过程模型，这项任务就简化为计算(14.109)中的期望．计算其闭型解或求数学期望的数值解都可以完成这项任务．如果存在闭型解，就可以积分得到 $E^Q[F(S_T,T)]$

$$\int_{S_{\min}}^{S_{\max}}F(S_T,T)\mathrm{d}Q(S_T) \tag{14.111}$$

246

其中 $\mathrm{d}Q(S_T)$是关于 S_T 发生的无穷小变化的鞅概率．$[S_{\min},\ S_{\max}]$是 S_T 可能的变化范围．在蒙特卡洛估值中，可以使用下面的近似：

$$E^Q[F(S_T,T)]\approx\frac{1}{N}\sum_{i=1}^{N}[F(S_T^j,T)] \tag{14.112}$$

其中 $j=1,\cdots,N$ 为指标，代表从无套利分布中随机选择的 S_T 轨迹．这种方法叫做蒙特卡洛方法．如果我们对测度 Q 下模型的随机成分刻画正确，且路径数 N 趋于无穷大，那么大数定律可以保证上面的平均值收敛于真实数学期望，因此这种近似效果很好．[㊀]

- 最后一步很简单．如果贴现因子 D_t 已知，除以 D_t 就得到现值．如果 D_t 随机，那么它的随机运动需要与 S_T 一起在期望运算符内进行考虑．

显然，Q 概率测度在这些计算中起到了非常重要的作用．由于使用了鞅概率，在定价过程中，我们不需要对 S_t 在真实概率分布 P 中建模，也不需要对风险溢价进行建模——这两种建模都很困难而且带有比较强的“主观性”．[㊁]

这带来了另一个问题．鞅概率对于金融从业者来说是一种重要工具，那么它对于计量经济学家而言也同样重要吗？

通常来说，这样的测度变换对计量经济学家的工作毫无帮助．这是因为计量经济学家们的目标是得到对 S_T 的最佳预测．那么使用 Q 测度会得到很糟的结果．为说明这一点，我们假设 S_T 在 T 时刻有 M 种可能状态．对 S_T 最好的预测记为$\widetilde{S}_T$，为：

$$\widetilde{S}_T=P_1S_T^1+\cdots+P_MS_T^M \tag{14.113}$$

㊀ 但这种方法也可能需要耗费大量时间和技术来得到合意的数值．

㊁ 相反，Q 测度是唯一的、客观的．所有从业者都在这一点上达成了共识．

$$=\sum_{i=1}^{M} P_j S_T^j \tag{14.114}$$

即通过将可能值乘以相应真实概率 P_j 求和得到 $\widetilde{S}_T$. [1] 显然，如果使用 Q 测度代替 P_j，得到的预测值

$$\widetilde{S}_T = Q_1 S_T^1 + \cdots + Q_M S_T^M \tag{14.115}$$

$$=\sum_{i=1}^{M} Q_j S_T^j \tag{14.116}$$

不能正确地反映 S_t 在区间 Δ 内的变化．这是因为 S_t 在 Q 测度下是以速率 $r_t\Delta$ 增长，而在真实测度下的期望增长率($(r_t+\mu)\Delta$)是包含风险溢价 μ 的．既然在鞅概率 Q 下对 S_t 的增量的表达是错误的，我们当然也无法得到令人满意的关于 S_T 预测值．至此，我们看出了虽然 Q 在定价过程中非常有用，但是当需要预测 S_T 时，我们还是应该使用真实概率 P 及 $E^P[\cdot]$.

14.7 如何得到等价概率

Girsanov 定理给出了一种利用鞅得到新概率的方法．例如，考虑一个非负随机过程 Z_t. 假设它对任意 t 在 P 概率下都具有下面性质：

247

$$E^P[Z_t]=1 \tag{14.117}$$

$$0\leqslant Z_t \tag{14.118}$$

我们将证明 Z_t 可以有效地帮助我们得到新概率．

考虑实数轴上的某个集合 A，定义其示性函数为 1_A：

$$1_A=\begin{cases}1 & 若\ Z_t\in A\\ 0 & 其他\end{cases} \tag{14.119}$$

此即，若 Z_t 的值落在 A 中则函数 1_A 取值 1，否则取值 0.

我们想知道下面这个表达式的含义：

$$E^P[Z_t 1_A] \tag{14.120}$$

其中 A 表示某个 Z_t 可能取值的集合．特别地，我们要证明该表达式为 Z_t 过程定义了一个新概率测度．

首先，Z_t 的期望值为 1. 事实上将该过程乘以指标函数 1_A 消除了落在 A 外的值．由于 Z_t 非负，我们一定有：

$$0\leqslant E^P[Z_t 1_A] \tag{14.121}$$

其次，假设 Ω 表示 Z_t 所有可能值，[2] 将其划分为 n 个互不相容的集合 A_i，使得

$$A_1\cup A_2\cup\cdots\cup A_n=\Omega \tag{14.122}$$

---

[1] 更准确地说，这里的 P_j 是条件概率．

[2] 如果 Z_t 表示金融资产价格，那么 Ω 为全部正实数．当存在最小报价单位时，Ω 是所有元素为正有理数的可数集．

这样，无论 Z_t 取何值，我们都有

$$1_{A_1}+\cdots+1_{A_n}=1_\Omega \tag{14.123}$$

于是

$$E^P[Z_t]=E^P[Z_t 1_\Omega] \tag{14.124}$$

或

$$E^P[Z_t]=E^P[Z_t 1_{A_1}]+E^P[Z_t 1_{A_2}]+\cdots+E^P[Z_t 1_{A_n}] \tag{14.125}$$

$$=1 \tag{14.126}$$

因此 $E^P[Z_t 1_{A_i}]$均非负，且和为 1.

如果将这些项写作

$$E^P[Z_t 1_{A_i}]=Q(A_i) \tag{14.127}$$

我们就得到了一个和 Z_t 相关的新概率：

$$Q(A_i)\geqslant 0 \tag{14.128}$$

及

$$\sum_{i=1}^{n}Q(A_i)=1 \tag{14.129}$$

这个新的概率测度 $Q(A_i)$和 Z_t 原始的真实概率 $P(A_i)$，

$$\Pr(Z_t\in A_i)=P(A_i) \tag{14.130}$$

之间差异可能很大.

因此在本例中，从真实概率测度 P 及其对应数学期望

$$E^P[Z_t 1_A] \tag{14.131}$$

出发，如果

$$E^P[Z_t]=1 \tag{14.132}$$

$$Z_t\geqslant 0 \tag{14.133}$$

我们可以得到一个新概率分布．如果过程 Z_t 为鞅，那么不同时间段内新概率集合也满足一致性. 248

不讨论技术性问题，我们直接来看表示这些概念的另一种表达方式．假设在真实概率 P 下，Z 的密度函数为 $f(z)$，那么

$$P(Z_t\in A_i)=\int_{A_i}f(z)\mathrm{d}z \tag{14.134}$$

那么根据定义有

$$E^P[Z_t]=\int_\Omega zf(z)\mathrm{d}z=1 \tag{14.135}$$

$$E^P[Z_t 1_A]=\int_\Omega 1_{A_i}zf(z)\mathrm{d}z \tag{14.136}$$

$$=\int_{A_i}zf(z)\mathrm{d}z=Q(A_i) \tag{14.137}$$

现在，假设我们要计算函数 $g(X_t)$在 P 概率测度下的数学期望：

$$E^P[g(X_t)] = \int_\Omega g(x)f(x)\mathrm{d}x \tag{14.138}$$

假设 $g(x)$像上面一样可以用 Z_t(依赖于 X_t)表示：

$$g(X_t) = Z_t h(X_t)$$

那么我们可以做下面这个变换：

$$E^P[g(X_t)] = \int_\Omega g(x)f(x)\mathrm{d}x \tag{14.139}$$

$$= \int_\Omega zh(x)f(x)\mathrm{d}x = \int_\Omega h(x)\tilde{f}(x)\mathrm{d}x = E^Q[h(X_t)] \tag{14.140}$$

有时我们会发现(14.140)中的积分要比(14.139)中的原始积分计算起来更简单．现在我们来看一个应用．

一个例子

考虑随机变量 Z_t，其定义为

$$Z_t = \mathrm{e}^{\left[\sigma W_t - \frac{1}{2}\sigma^2 t\right]} \tag{14.141}$$

其中 W_t 是 P 概率测度下的维纳过程，均值为 0，方差为 t. $\sigma>0$ 为已知常数．

根据定义，$Z_t \geqslant 0$. 我们首先考虑其一阶矩．直接求数学期望：

$$E^P[Z_t] = \int_{-\infty}^{\infty} \mathrm{e}^{\left[\sigma W_t - \frac{1}{2}\sigma^2 t\right]} \frac{1}{\sqrt{2\pi t}} \mathrm{e}^{-\frac{1}{2t}W_t^2} \mathrm{d}W_t \tag{14.142}$$

化简为

$$E^P[Z_t] = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}} \mathrm{e}^{-\frac{1}{2t}(W_t^2 - 2\sigma t W_t + \sigma^2 t^2)} \mathrm{d}W_t \tag{14.143}$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}} \mathrm{e}^{-\frac{1}{2t}(W_t - \sigma t)^2} \mathrm{d}W_t \tag{14.144}$$

注意到被积函数其实是均值为 σt、方差为 t 的正态随机变量的密度函数，所以我们一定有

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}} \mathrm{e}^{-\frac{1}{2t}(W_t - \sigma t)^2} \mathrm{d}W_t = 1 \tag{14.145}$$

这意味着

$$E^P[Z_t] = 1 \tag{14.146}$$

因此 $\mathrm{e}^{\left[\sigma W_t - \frac{1}{2}\sigma^2 t\right]}$满足前一节中讨论的非负过程 Z_t 的条件．由于它大于零且数学期望等于 1，我们可以用它来生成等价概率测度．

事实上，对于任何一个集合 $A \subseteq \mathbf{R}$，我们可以从初始概率 P 出发，通过计算期望：

249

$$Q(A) = E^P\left[\mathrm{e}^{\left[\sigma W_t - \frac{1}{2}\sigma^2 t\right]} 1_A\right] \tag{14.147}$$

定义新的概率测度．这个公式该如何应用于金融证券定价呢?

考虑执行价格为 K 的看涨期权无套利价值 C_t. 我们知道，C_t 在 Black-Scholes 假设下为：

$$C_t = \mathrm{e}^{-r(T-t)} E^P[\max\{S_T - K, 0\}] \tag{14.148}$$

在 Black-Scholes 的假设下，S_t 在风险中性概率测度 Q 下是几何过程，服从 SDE：

$$\mathrm{d}S_t = rS_t\mathrm{d}t + \sigma S_t \mathrm{d}W_t \tag{14.149}$$

我们已知这个 SDE 的一个解为：

$$S_t = S_0 \mathrm{e}^{\left[rt + \sigma W_t - \frac{1}{2}\sigma^2 t\right]} \tag{14.150}$$

用它代换(14.148)中 S_T，(14.148)中期权价格为：

$$C_t = \mathrm{e}^{-r(T-t)} E^P\left[\max\left[S_t \mathrm{e}^{r(T-t)+\sigma(W_T - W_t) - \frac{1}{2}\sigma^2(T-t)} - K, 0\right]\right] \tag{14.151}$$

很有趣的一点在于，我们可以把(14.141)定义的变量 Z_t 以某种形式嵌套入在这个表达式中．具体来说，我们将该指数项分成两部分，可以写成：

$$E^P\left[\max\left[S_t \mathrm{e}^{r(T-t)+\sigma(W_T - W_t) - \frac{1}{2}\sigma^2(T-t)} - K, 0\right]\right] \tag{14.152}$$

$$= E^P\left[\max\left[S_t \mathrm{e}^{\sigma(W_T - W_t) + (r - \frac{\sigma^2}{2})(T-t)} - K, 0\right]\right] \tag{14.153}$$

提出因子后：

$$E^P\left[\mathrm{e}^{\sigma(W_T - W_t) - \frac{1}{2}\sigma^2(T-t)} \max\left[S_t \mathrm{e}^{r(T-t)} - \mathrm{e}^{-\sigma(W_T - W_t) + \frac{1}{2}\sigma^2(T-t)} K, 0\right]\right] \tag{14.154}$$

与前面相同，令

$$Z_T = \mathrm{e}^{\sigma(W_T - W_t) - \frac{1}{2}\sigma^{2(T-t)}}$$

我们可以得到

$$E^P\left[\mathrm{e}^{\sigma(W_T - W_t) - \frac{1}{2}\sigma^2(T-t)} \max\left[S_t \mathrm{e}^{r(T-t)} - \mathrm{e}^{-\sigma(W_T - W_t) + \frac{1}{2}\sigma^2(T-t)} K, 0\right]\right]$$

$$= E^P\left[Z_T \max\left[S_t \mathrm{e}^{r(T-t)} - \mathrm{e}^{-\sigma(W_T - W_t) + \frac{1}{2}\sigma^2(T-t)} K, 0\right]\right] \tag{14.155}$$

$$= E^Q\left[\max\left[S_t \mathrm{e}^{r(T-t)} - \mathrm{e}^{-\sigma(W_T - W_t) + \frac{1}{2}\sigma^2(T-t)} K, 0\right]\right] \tag{14.156}$$

其中 Q 概率测度定义为：

$$Q(A) = E^P[Z_t 1_A]$$

通过变换到 Q 概率测度，表示 Z_T 的项消失了，而其数学期望更容易计算．当对奇异期权进行定价时，使用这种方法的变换很易得到定价公式．从本质上讲，我们可以看出关于几何过程的数学期望都会含有由 Z_t 代表的隐含项．本节讨论的技巧使变换测度变为可能，而得到的新测度下的数学期望更容易计算．

14.8 小结

在此基础上，我们来回顾一下将 S_t 变为鞅过程的重要步骤．

- 通过将 S_t 的分布从 P 测度变换到 Q 测度，引入新误差项 $\widetilde{W}_t$，可以完成该变换．
- 新误差项 $\widetilde{W}_t$ 和原始误差项方差一致．
- (14.108)与(14.94)的差别在于 S_t 的均值发生了变化，而误差项保持 0 均值不变． 250

我们是通过测度变换，而不是直接从随机变量中减去常数得到的．

- 更重要的是，在本例中，该变换用于将 S_t 变为鞅．在金融模型中，我们可能对 $e^{-rt}S_t$ 做测度变换而不是对 S_t 做变换．$e^{-rt}S_t$ 代表资产价格按（无风险）利率 r 进行贴现后的现值．此时，我们需要对函数 $\xi(S_t)$ 重新定义．

14.9 参考阅读

通过 Girsanov 定理将随机过程变换为鞅，是随机微积分中比较深刻的主题．为这个方法提供技术背景的资料都比较艰深．Karatzas 和 Shreve(1991)给出了一个更直观的讨论. Liptser 和 Shiryayev(1977)则是一本比较全面的参考读物．

14.10 习题

1. 考虑随机变量 Δx，它的取值及相应概率如下：

$$\{\Delta x = 1, p(\Delta x = 1) = 0.3\}$$
$$\{\Delta x = -0.5, p(\Delta x = -0.5) = 0.2\}$$
$$\{\Delta x = 0.2, p(\Delta x = 0.2) = 0.5\}$$

(a) 计算该随机变量的均值和方差．

(b) 将 Δx 减去适当的常数，使该随机变量的均值变为 0.05. 即计算

$$\Delta y = \Delta x - \mu$$

使 Δy 均值为 0.05.

(c) 方差是否发生变化？

(d) 改变概率测度，仍保持方差不变，进行相同的变换．

(e) Δx 的值是否发生变化？

2. 假设股票收益率 R_t 在 t 固定时服从对数正态分布：

$$\log(R_t) \sim \mathcal{N}(\mu, \sigma^2)$$

将 $\log(R_t)$ 的密度函数记为 $f(R_t)$，假设 $\mu=0.17$，进一步估计方差为 $\sigma^2=0.09$.

(a) 求函数 $\xi(R_t)$，使得 R_t 在密度函数 $\xi(R_t)f(R_t)$ 下均值等于无风险利率 $r=0.05$.

(b) 求函数 $\xi(R_t)$，使得 R_t 均值为 0.

(c) 在哪种概率测度下计算

$$E[R_t^2]$$

更容易？

(d) 在这两个概率测度下的方差是否不同？

3. 已知长期利率 R 和短期利率 r 服从二元正态分布，方差-协方差矩阵为 Σ，均值为 μ，

$$\Sigma = \begin{bmatrix} 0.5 & 0.1 \\ 0.1 & 0.9 \end{bmatrix} \quad \mu = \begin{bmatrix} 0.07 \\ 0.05 \end{bmatrix}$$

相应密度函数记为 $f(R, r)$.

(a) 用 Mathematica 或 Maple 画出该密度函数的图像．

(b) 求函数 $\xi(R, r)$，使利率在概率测度：

$$dQ = \xi(R, r) f(R, r) dR dr$$

下均值为 0.

(c) 画出 $\xi(R,r)$ 及新密度函数的图像.

(d) 利率的方差-协方差矩阵是否发生变化?

4. 利用蒙特卡洛算法估计标准正态分布的样本小于-4的概率. 令 $N=1000$，随机变量均值为 0，方差为 1. 多次运行该算法，观察估计值的变化程度. 对这种算法的有效性和稳定性做出评价. 接下来使用 Girsanov 定理改进该估计. 描述所使用的变换. 估计值得到改进了吗?

5. 考虑以下欧式障碍期权. 在未来的 30 天内，如果当其标的股票价格超过 110 美元时，这个期权的拥有者可以 100 美元的价格购入股票. 假设股票价格服从几何布朗运动，其中 $r=0.05$，$\sigma=0.2$，估计该期权价格.

252

第 15 章　等价鞅测度

15.1　引言

在本章中，我们将利用期权定价说明如何应用等价鞅测度．对于以股票 S_t 为标的证券（不支付分红）的欧式看涨期权 C_t，计算它的无套利价格有两种方法．

1）原始 Black-Scholes 方法：①构建无风险投资组合；②得到关于 $F(S_t,t)$ 的偏微分方程；③直接求解这个 PDE，或者用数值方法解该 PDE 的近似．

2）鞅方法，找到一个新的概率测度 Q 与原测度等价，使得 S_t 在 Q 下变为鞅，然后再使用分析方法或数值方法计算数学期望

$$C_t = E^Q e^{r(T-t)}[\max\{S_T - K, 0\}] \tag{15.1}$$

本章第一部分逐步介绍鞅方法．我们从 Black 和 Scholes 给出的假设出发，说明如何将（贴现后的）资产价格变为鞅．我们可以通过找到等价鞅测度 Q 做到这一点，此处不会直接用到 Girsanov 定理．

253

Girsanov 定理将用于本章的第二部分，讨论两种资产定价方法之间的联系．特别地，我们将说明，将（贴现后的）看涨期权价格变为鞅等价于令 $F(S_t,t)$ 满足前面的 Black-Scholes 偏微分方程．最终我们会发现 PDE 方法与鞅方法密切相关．

15.2　鞅测度

第 12 章中讨论过无风险投资组合的构建方法和 PDE 的使用，但是在那里我们并没有给出 Black-Scholes 公式的逐步推导．

等价鞅测度方法采用不同的方式得到相同的公式．公式的推导可能略显冗长乏味，但它背后的数学知识却很清晰．我们将用这种方法一步步给出 Black-Scholes 公式的推导．

先来讨论几个中间结果．这些结果本身都很重要，因为在资产定价公式中会频繁用到它们.

15.2.1　矩母函数

Y_t 为连续时间随机过程，㊀

$$Y_t \sim \mathcal{N}(\mu t, \sigma^2 t) \tag{15.2}$$

其中 Y_0 已知．

定义 S_t 为几何布朗运动

㊀ 有时也称 Y_t 为广义维纳过程，因为它服从正态分布，但是均值非零，方差不一定等于 1.

$$S_t = S_0 e^{Y_t} \tag{15.3}$$

S_0 是 S_t 的起始点，值是外生给定的．[⊖]我们首先介绍 Y_t 的矩母函数．

将该矩母函数记为 $M(\lambda)$，它是关于 Y_t 的一个函数的数学期望，

$$M[\lambda] = E[e^{\lambda Y_t}] \tag{15.4}$$

其中 λ 为任意参数．矩母函数在资产定价公式中非常有用．此外，计算矩母函数的过程将会涉及一些随机积分中的标准运算．单就这一点而言，我们也需要详细阐述这个计算过程．

矩母函数的计算

利用(15.2)中的分布，我们可以计算出 $E[e^{rY_t}]$. 根据定义，(15.4)可以写成

$$E[e^{\lambda Y_t}] = \int_{-\infty}^{\infty} e^{\lambda Y_t} \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-\frac{1}{2}\frac{(Y_t-\mu t)^2}{\sigma^2 t}} dY_t \tag{15.5}$$

整理指数项，可得：

$$E[e^{\lambda Y_t}] = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-\frac{1}{2}\frac{(Y_t-\mu t)^2}{\sigma^2 t}+\lambda Y_t} dY_t \tag{15.6}$$

此时被积函数的指数项并不是一个完全平方式，但可以通过在右边乘以

$$e^{-(\lambda\mu t+\frac{1}{2}\sigma^2 t\lambda^2)} e^{(\lambda\mu t+\frac{1}{2}\sigma^2 t\lambda^2)} = 1 \tag{15.7}$$

254

化为完全平方式．则(15.6)中的等式变为

$$E[e^{\lambda Y_t}] = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{(\lambda\mu t+\frac{1}{2}\sigma^2 t\lambda^2)} \times e^{-\frac{1}{2}\frac{(Y_t-\mu t)^2}{\sigma^2 t}+\lambda Y_t-(\lambda\mu t+\frac{1}{2}\sigma^2 t\lambda^2)} dY_t \tag{15.8}$$

现在第二个指数函数的指数可以凑成平方式．该项与 Y_t 无关，可以移到积分号外．这样我们得到

$$E[e^{\lambda Y_t}] = e^{(\lambda\mu t+\frac{1}{2}\sigma^2\lambda^2 t)} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-\frac{1}{2}\frac{(Y_t-(\mu t+\sigma^2 t\lambda))^2}{\sigma^2 t}} dY_t \tag{15.9}$$

而该表达式右边的被积函数是正态分布随机变量的密度函数，因此其积分等于 1. 我们得到

$$M(\lambda) = e^{(\lambda\mu t+\frac{1}{2}\sigma^2\lambda^2 t)} \tag{15.10}$$

矩母函数在统计学中是很有用的工具．计算它关于 λ 的 k 阶导数，并令 $\lambda=0$，可以得到随机变量的 k 阶矩．

例如，计算(15.10)关于 λ 的导数可以得到 Y_t 的一阶矩：

$$\frac{\partial M}{\partial\lambda} = (\mu t + \sigma^2 t\lambda) e^{(\lambda\mu t+\frac{1}{2}\sigma^2\lambda^2 t)} \tag{15.11}$$

将 $\lambda=0$ 代入该公式，得到

$$\left.\frac{\partial M}{\partial\lambda}\right|_{\lambda=0} = \mu t \tag{15.12}$$

⊖ S_0 可以是随机的，只要它独立于 Y_t.

如果要得到二阶矩，只需要求二阶导数，并令 $\lambda=0$：

$$\left.\frac{\partial^2 M}{\partial\lambda^2}\right|_{\lambda=0}=\sigma^2 t \tag{15.13}$$

这些都是非常有用的性质，但它们在资产定价问题中起到的作用不是太重要．矩母函数在资产定价中的实用性体现在(15.10)，或者直接把这个结果写为

$$E[e^{\lambda Y_t}]=e^{\left(\lambda\mu t+\frac{1}{2}\sigma^2\lambda^2 t\right)} \tag{15.14}$$

后面我们会看到，需要求几何布朗运动的数学期望时，这个结果会使计算非常方便——因为它直接给出了几何布朗运动数学期望的确切值．

15.2.2 几何布朗运动的条件期望

当用鞅方法为金融衍生品定价时，需要计算条件期望 $E[S_t|S_u,u<t]$，其中 S_t 是之前讨论过的几何布朗运动．这是我们在讨论鞅方法前要讨论的第二个中间结果．假设与前一节相同，

$$S_t=S_0e^{Y_t},\quad t\in[0,\infty) \tag{15.15}$$

其中 Y_t 仍服从分布

$$Y_t\sim\mathcal{N}(\mu t,\sigma^2 t) \tag{15.16}$$

根据定义，总有

$$Y_t=Y_s+\int_s^t dY_u \tag{15.17}$$

定义 ΔY_t 为

$$\Delta Y_t=\int_s^t dY_u \tag{15.18}$$

根据广义维纳过程的定义，我们有

255

$$\Delta Y_t\sim\mathcal{N}(\mu(t-s),\sigma^2(t-s)) \tag{15.19}$$

因此 ΔY_t 也是正态分布随机变量．根据前一节可知其矩母函数为

$$M(\lambda)=e^{\lambda\mu(t-s)+\frac{1}{2}\sigma^2\lambda^2(t-s)} \tag{15.20}$$

利用这些结果，我们可以计算出几何布朗运动的条件期望．首先，由于 S_u 在 u 时刻是非随机的，我们有

$$E\left[\frac{S_t}{S_u}\,|\,S_u,u<t\right]=E[e^{\Delta Y_t}\,|\,S_u] \tag{15.21}$$

因为当 $u<t$ 时，ΔY_t 与 Y_u 相互独立，所以

$$E[e^{\Delta Y_t}\,|\,S_u]=E[e^{\Delta Y_t}] \tag{15.22}$$

此外，$E[e^{\Delta Y_t}]$是(15.20)中的矩母函数在 $\lambda=1$ 处的函数值，因此

$$E[e^{\Delta Y_t}]=e^{\mu(t-u)+\frac{1}{2}\sigma^2(t-u)} \tag{15.23}$$

$$=E\left[\frac{S_t}{S_u}\,|\,S_u\right] \tag{15.24}$$

或者两边同时乘以 S_u，

$$E[S_t \mid S_u] = S_u e^{\mu(t-u)+\frac{1}{2}\sigma^2(t-u)} \tag{15.25}$$

该公式给出了几何布朗运动的条件期望．这个结果在资产定价理论经常得到使用，下面的讨论中也会涉及．

15.3 将资产价格转化为鞅

和上一节一样，我们假设

$$S_t = S_0 e^{Y_t}, \quad t \in [0,\infty) \tag{15.26}$$

其中 Y_t 是维纳过程，其分布记为 P. 那么 P 是影响资产价格 S_t 的无穷小波动所在真实测度．

理论上，S_t 的观察值所对应的分布是 P 测度下的概率分布．但对金融分析师而言，这个分布下的计算并不方便．根据第 14 章中的讨论，也许我们可以找到一个等价概率测度 Q，使资产定价过程中的计算更容易．如果资产价格过程在这个新的概率测度下是鞅，那么计算方便这一优点尤其明显．

我们在本节中利用一个例子讨论如何得到这种概率测度．

S_t 的真实分布由 Y_t 的分布决定．在 P 概率测度下，Y_t 的概率分布服从

$$Y_t \sim \mathcal{N}(\mu t, \sigma^2 t) \tag{15.27}$$

假设 S_t 表示标的资产在 t 时刻的价格，并用 $S_u(u<t)$ 表示在稍早的 u 时刻观察到的资产价格.

首先，由于资产 S_t 是有风险的，因此当它按无风险利率贴现时，得到的结果一定不是鞅．也就是说，在真实概率测度 P 下，我们得不到鞅等式

$$E^P[e^{-rt}S_t \mid S_u, u < t] = e^{-ru}S_u \tag{15.28}$$

实际上，由于其资产的价格中包含风险溢价，我们得到的通常是

$$E^P[e^{-rt}S_t \mid S_u, u < t] > e^{-ru}S_u \tag{15.29}$$

所以，在真实概率测度 P 下，下面定义的这个随机过程 Z_t 不可能是鞅．

$$Z_t = e^{-rt}S_t \tag{15.30}$$

但我们还是可以用第 14 章中所介绍的方法来改变 Z_t 的漂移项，使它变为鞅：在一定的条件下，我们可以得到等价鞅测度 Q，使得下面这个鞅等式成立

$$E^Q[e^{-rt}S_t \mid S_u, u < t] = e^{-ru}S_u \tag{15.31}$$

256

也可以将该等式用 Z_t 表示：

$$E^Q[Z_t \mid Z_u, u < t] = Z_u \tag{15.32}$$

正如我们在第 14 章中所看到的，我们可以把误差项中的维纳过程 W_t 变换成 Q 测度下新随机过程 W_t^*，而变换后的 dZ_t 的漂移项为零．

我们在下一节中讨论如何找到这样的概率测度变换．

15.3.1 确定测度 Q

现在我们的问题是如何找到概率测度 Q，使得该测度下的条件期望满足鞅等式

$$E^{Q}[\mathrm{e}^{-rt}S_t \mid S_u, u<t] = \mathrm{e}^{-ru}S_u \tag{15.33}$$

这代表 S_t 是测度 Q 下的鞅．[⊖]

如何确定这样的 Q 呢？它的形式是怎样的呢？

接下来我们一步一步地推导出这个测度．已知

$$S_t = S_0\mathrm{e}^{Y_t} \tag{15.34}$$

其中 Y_t 在概率测度 P 下的分布为：

$$Y_t \sim \mathcal{N}(\mu t, \sigma^2 t) \tag{15.35}$$

定义它在概率测度 Q 下的分布为：

$$Y_t \sim \mathcal{N}(\rho t, \sigma^2 t) \tag{15.36}$$

其中漂移参数 ρ 可任取，这是该分布在两种测度下唯一的差异——这两个概率分布具有相同的方差．

我们可以利用(15.36)求出条件期望

$$E^{Q}[\mathrm{e}^{-rt}S_t \mid S_u, u<t] \tag{15.37}$$

我们只需使用(15.25)，就直接可以得到

$$E^{Q}[\mathrm{e}^{-r(t-u)}S_t \mid S_u, u<t] = S_u\mathrm{e}^{-r(t-u)}\mathrm{e}^{\rho(t-u)+\frac{1}{2}\sigma^2(t-u)} \tag{15.38}$$

由于这是在概率测度 Q 下计算的数学期望，公式右边与 ρ 相关，与 μ 无关．

由于(15.36)中参数 ρ 可任取，我们为它选择一个合意的值，使概率测度 Q 下的数学期望满足鞅条件．为此，我们令

$$\rho = r - \frac{1}{2}\sigma^2 \tag{15.39}$$

可以看到，这里参数 ρ 是由波动率 σ 和无风险利率 r 确定．这样的 ρ 可以令

$$-r(t-u) + \rho(t-u) + \frac{1}{2}\sigma^2(t-u) = 0 \tag{15.40}$$

因而(15.38)右边的指数部分等于 1. 将这个结果代入(15.38)，可得

$$E^{Q}[\mathrm{e}^{-r(t-u)}S_t \mid S_u, u<t] = S_u \tag{15.41}$$

将 e^{ru} 移到右边，

$$E^{Q}[\mathrm{e}^{-rt}S_t \mid S_u, u<t] = S_u\mathrm{e}^{-ru} \tag{15.42}$$

我们得到了一个鞅等式，这说明 $\mathrm{e}^{-rt}S_t$ 是在概率测度 Q 下的鞅过程．

我们可以通过确定 ρ 的值找到一个概率分布，使资产价格在这个概率分布下的数学期望满足鞅条件．在本节讨论的这个例子里，Y_t 在新测度下服从如下的正态分布

$$\mathcal{N}\left(\left(r-\frac{1}{2}\sigma^2\right)t, \sigma^2 t\right) \tag{15.43}$$

257 该概率分布不同于(15.35)中真实概率测度 P 下的分布(它们的均值不同).

⊖ 假设 S_t 满足作为鞅的其他条件要求，比如可积性．

15.3.2 隐含 SDE

前一节讨论了如何确定等价鞅测度 Q. 我们在这一节中比较在这两个等价概率测度下，资产价格过程所满足的随机微分方程(SDE). 这对于我们进一步了解这种等价鞅测度变换的意义很有帮助.

和上一节一样，假设 S_t 的定义如下

$$S_t = S_0 e^{Y_t}, \quad t \in [0, \infty) \tag{15.44}$$

其中 Y_t 服从均值为 μt、方差为 $\sigma^2 t$ 的正态分布(在"真实"概率测度 P 下). 也就是说，增量 dY_t 满足

$$dY_t = \mu dt + \sigma dW_t \tag{15.45}$$

要得到 S_t 满足的 SDE，我们需要先得到随机微分 dS_t 的表达式. 由于 S_t 是 Y_t 的函数，且后者满足的 SDE 已知，因此可以直接使用伊藤引理得到：

$$dS_t = S_0 e^{Y_t}[\mu dt + \sigma dW_t] + [S_0 e^{Y_t}] \frac{1}{2}\sigma^2 dt \tag{15.46}$$

移项，代入 S_t，(15.46)可以化简为

$$dS_t = \left[\mu S_t + \frac{1}{2}\sigma^2 S_t\right]dt + \sigma S_t dW_t \tag{15.47}$$

所以，在真实测度 P 下，关于描述资产价格 S_t 的 SDE 模型中，我们有下列结论：

1) 漂移系数为 $\left(\mu + \frac{1}{2}\sigma^2\right)S_t$；

2) 扩散系数为 σS_t；

3) 驱动过程为测度 P 下的维纳过程 W_t.

在鞅测度 Q 下 SDE 的计算方法类似，但漂移系数不同. 只需替换(15.47)中的 W_t 和 μ(变为 ρ)就可以得到该 SDE. 经过类似的步骤后，我们可以得到

$$dS_t = \left[\rho S_t + \frac{1}{2}\sigma^2 S_t\right]dt + \sigma S_t dW_t^* \tag{15.48}$$

这里我们要强调的是一个之前没有注意到的重要步骤：替换 W_t 意味着我们将概率测度从 P 变为 Q. 因为只有在 Q 测度下，等式(15.48)中的误差项才是标准维纳过程. 如果继续使用 P 测度，误差项本身就会含有一个非零漂移项.

现在，我们可以令等式(15.48)中 ρ 等于

$$\rho = r - \frac{1}{2}\sigma^2 \tag{15.49}$$

将它代入(15.48)，可得

$$dS_t = \left[\left(r - \frac{1}{2}\sigma^2\right)S_t + \frac{1}{2}\sigma^2 S_t\right]dt + \sigma S_t dW_t^* \tag{15.50}$$

含有 $\frac{1}{2}\sigma^2$ 的项相互抵消，得到 SDE：

$$dS_t = rS_t\,dt + \sigma S_t\,dW_t^* \tag{15.51}$$

这个是一个很有趣的结果：Q 测度将 S_t 变为鞅的同时，将原始 SDE 中漂移系数变为无风险利率 r. 原始模型中 μ 包含风险溢价，算出 S_t 前它一般是未知的，而无风险利率 r 是由已知给出的．

两个 SDE 中还有第二点不同．(15.51)中的 SDE 由新维纳过程 W_t^* 驱动，具有测度 Q 下的分布，它与真实世界中出现各种状态的可能性(由 P 概率测度确定)都无关．除此以外，这种测度计算起来很方便．(贴现后的)资产价格在该测度下是鞅，这一点在对衍生资产进行定价时是很好的性质．并且金融理论说明，在不存在套利空间时，这种使资产价格变为鞅的合成概率 Q 一定存在．

258

15.4 应用：Black-Scholes 公式

当满足下列条件时，Black-Scholes 公式给出了看涨期权的价格 $F(S_t,t)$：

1) 无风险利率在期权有效期内为常数；
2) 标的证券在期权到期前不支付股息；
3) 该看涨期权是欧式期权，不可在到期日前行权；
4) 标的证券的价格 S_t 是几何布朗运动，它的漂移项和扩散项均与 S_t 成比例；
5) 不存在交易费，且资产可无限分割．

在这些条件下，下面这个 PDE 的解析解就是 Black-Scholes 公式：

$$0 = -rF + F_t + rF_sS_t + \frac{1}{2}\sigma^2 F_{ss}S_t^2, \quad 0 \leqslant S_t, \quad 0 \leqslant t \leqslant T \tag{15.52}$$

其中边界条件为 $F(S_t,t)=\max[S_T-K,0]$.

而这个解析解就是

$$F(S_t,t) = S_tN(d_1) - Ke^{-r(T-t)}N(d_1 - \sigma\sqrt{T-t}) \tag{15.53}$$

其中

$$d_1 = \frac{\log\left(\frac{S_t}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}} \tag{15.54}$$

在该表达式中，T 表示看涨期权的到期日，r 表示无风险利率，K 表示执行价格，σ 表示波动率，函数 $N(x)$表示标准正态分布随机变量小于 x 的概率．例如，$N(d_1)$等于

$$N(d_1) = \int_{-\infty}^{d_1} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,dx \tag{15.55}$$

令 S_t 表示标的资产，C_t 表示以该资产为标的的欧式看涨期权价格．在标准 Black-Scholes 条件下，不发生分红，无风险利率为常数，无交易费．我们接下来要做的是利用等价鞅测度 Q 来直接推导出 Black-Scholes 公式．

在测度 Q 下，$e^{-rt}C_t$ 应满足的最基本的关系应该是鞅性质

$$C_t = E^Q[e^{-r(T-t)}C_T] \tag{15.56}$$

其中 $T>t$ 为看涨期权到期日.

到期时，若 $S_T>K$，则期权收益为 S_T-K，否则该看涨期权到期时价值为 0. 因此边界条件为

$$C_T = \max[S_T - K, 0] \tag{15.57}$$

$e^{-rt}C_t$ 的鞅性质变为

$$C_t = E_t^Q[e^{-r(T-t)}\max\{S_T - K, 0\}] \tag{15.58}$$

导出 Black-Scholes 公式需要确切计算出该期望. 推导过程很简单，但表达式很长，最好对记法做出简化. 我们做出如下简化:

- 令 $t=0$，计算期权在零时刻的价格.
- 相应地，信息集 I_t 变为 I_0. 这样我们可以使用无条件期望运算符 $E^Q[\cdot]$，而不是计算条件期望.

我们现在开始逐步推导 Black-Scholes 公式: 在概率测度 Q 下直接计算

$$C_0 = E^Q[e^{-rT}\max\{S_T - K, 0\}] \tag{15.59}$$

概率测度 Q 为等价鞅测度，在前一节已求出

$$dQ = \frac{1}{\sqrt{2\pi\sigma^2 T}} e^{-\frac{1}{2\pi\sigma^2 T}(Y_T - (r - \frac{1}{2}\sigma^2)T)^2} dY_T \tag{15.60}$$

259

其中

$$S_T = S_0 e^{Y_T} \tag{15.61}$$

利用这个密度函数，我们可以直接计算

$$C_0 = E^Q[e^{-rT}\max\{S_T - K, 0\}] \tag{15.62}$$

的值，它也可以写成

$$C_0 = \int_{-\infty}^{\infty} e^{-rT}\max\{S_T - K, 0\}]dQ \tag{15.63}$$

其中仍有

$$S_T = S_0 e^{Y_T} \tag{15.64}$$

将它们带入(15.63)，有

$$C_0 = \int_{-\infty}^{\infty} e^{-rT}\max[S_0 e^{Y_T} - K, 0] \times \frac{1}{\sqrt{2\pi\sigma^2 T}} e^{-\frac{1}{2\pi\sigma^2 T}(Y_T - (r - \frac{1}{2}\sigma^2)T)^2} dY_T \tag{15.65}$$

可以改变积分下限来去掉积分内的最大值函数. 取对数后,

$$S_0 e^{Y_T} \geqslant K \tag{15.66}$$

等价于

$$Y_T \geqslant \log\left(\frac{K}{S_0}\right) \tag{15.67}$$

代入(15.65)

$$C_0 = \int_{\log(\frac{K}{S_0})}^{\infty} e^{-rT}(S_0 e^{Y_T} - K) \frac{1}{\sqrt{2\pi\sigma^2 T}} e^{-\frac{1}{2\pi\sigma^2 T}(Y_T - (r - \frac{1}{2}\sigma^2)T)^2} dY_T \tag{15.68}$$

该积分可以被分成两部分：

$$C_0 = S_0\int_{\log\left(\frac{K}{S_0}\right)}^{\infty} \mathrm{e}^{-rT}\mathrm{e}^{Y_T}\frac{1}{\sqrt{2\pi\sigma^2 T}}\mathrm{e}^{-\frac{1}{2\pi\sigma^2 T}\left(Y_T-\left(r-\frac{1}{2}\sigma^2\right)T\right)^2}\mathrm{d}Y_T + K\mathrm{e}^{-rT}\int_{\log\left(\frac{K}{S_0}\right)}^{\infty}\frac{1}{\sqrt{2\pi\sigma^2 T}}\mathrm{e}^{-\frac{1}{2\pi\sigma^2 T}\left(Y_T-\left(r-\frac{1}{2}\sigma^2\right)T\right)^2}\mathrm{d}Y_T \tag{15.69}$$

接下来我们分别计算表达式右边的两个积分.

计算

首先，为了进一步简化记法，我们定义一个新随机变量 Z

$$Z = \frac{Y_T - \left(r-\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}} \tag{15.70}$$

这使积分下限发生了变化，(15.69)右边第二个积分项变为

$$K\mathrm{e}^{-rt}\int_{\log\left(\frac{K}{S_0}\right)}^{\infty}\frac{1}{\sqrt{2\pi\sigma^2 T}}\mathrm{e}^{-\frac{1}{2\pi\sigma^2 T}\left(Y_T-\left(r-\frac{1}{2}\sigma^2\right)T\right)^2}\mathrm{d}Y_T = K\mathrm{e}^{-rt}\int_{\frac{\log\left(\frac{K}{S_0}\right)-\left(r-\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}}^{\infty}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}Z^2}\mathrm{d}Z \tag{15.71}$$

该积分下限与 Black-Scholes 公式中参数 d_2 密切相关. ⊖ 令

260

$$-\log\left(\frac{K}{S_0}\right) = \log\left(\frac{S_0}{K}\right) \tag{15.73}$$

得到 Black-Scholes 公式中的参数 d_2：

$$-\frac{\log\left(\frac{S_0}{K}\right)+\left(r-\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}} = -d_2 \tag{15.74}$$

由于正态分布具有对称性，当 $f(x)$表示标准正态分布密度函数时，有

$$\int_L^{\infty} f(x)\mathrm{d}x = \int_{-\infty}^{-L} f(x)\mathrm{d}x \tag{15.75}$$

根据(15.74)和(15.75)，有

$$K\mathrm{e}^{-rT}\int_{-d_2}^{\infty}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}Z^2}\mathrm{d}Y_T = K\mathrm{e}^{-rT}\int_{-\infty}^{d_2}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}Z^2}\mathrm{d}Y_T \tag{15.76}$$

$$= K\mathrm{e}^{-rT}N(d_2) \tag{15.77}$$

因此，我们得到了 Black-Scholes 公式的第二部分及参数 d_2 的值.

⊖ 积分下限变化的原因在于，当 Y_T 在 $\log\left(\frac{K}{S_0}\right)$ 与∞之间变化时，(15.70)定义的 Z 将在

$$\frac{\log\left(\frac{S_0}{K}\right)-\left(r-\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}} \tag{15.72}$$

与无穷大之间变化.

下面推导第一部分 $S_0 N(d_1)$，并说明 d_1 与 d_2 之间的关系．这就要对(15.69)右边第一个积分项进行变换．与之前相同，仍使用(15.70)定义的变量 Z：

$$\int_{\log\left(\frac{K}{S_0}\right)}^{\infty} \mathrm{e}^{-rT} S_0 \mathrm{e}^{Y_T} \frac{1}{\sqrt{2\pi\sigma^2 T}} \mathrm{e}^{-\frac{1}{2\pi\sigma^2 T}\left(Y_T-\left(r-\frac{1}{2}\sigma^2\right)T\right)^2} \mathrm{d}Y_T$$
$$= \mathrm{e}^{\left(r-\frac{1}{2}\sigma^2\right)} \mathrm{e}^{-rT} S_0 \int_{\frac{\log\left(\frac{K}{S_0}\right)-\left(r-\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}}^{\infty} \mathrm{e}^{\sigma Z\sqrt{T}} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}Z^2} \mathrm{d}Z \tag{15.78}$$

利用正态分布密度函数的对称，对右边的积分进行变换：

$$= \mathrm{e}^{-rT} S_0 \mathrm{e}^{\left(r-\frac{1}{2}\sigma^2\right)T} \int_{-\infty}^{d_2} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\left(Z^2+2\sigma Z\sqrt{T}\right)} \mathrm{d}Z \tag{15.79}$$

下面，在指数中加上再减去

$$\frac{\sigma^2 T}{2} \tag{15.80}$$

凑成完全平方，我们就可以得到

$$= \mathrm{e}^{-rT} S_0 \mathrm{e}^{\frac{T\sigma^2}{2}} \mathrm{e}^{\left(r-\frac{1}{2\sigma^2}\right)T} \int_{-\infty}^{d_2} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\left(Z-\sigma\sqrt{T}\right)^2} \mathrm{d}Z \tag{15.81}$$

积分号前除 S_0 外的项都互相抵消．最后，做如下的变量替换

$$H = Z + \sigma\sqrt{T} \tag{15.82}$$

我们可以得到

$$= S_0 \int_{-\infty}^{d_2+\sigma\sqrt{T}} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}H^2} \mathrm{d}H \tag{15.83}$$

其中

$$d_1 = d_2 + \sigma\sqrt{T} \tag{15.84}$$

该公式给出了 Black-Scholes 公式的第一部分，推导完毕．注意，在整个该推导过程中我们并没有去求解 PDE，却可以在最后得到 PDE 的解析解．

15.5 鞅方法与 PDE 方法的比较

我们已经看到，两种方法都可以计算出衍生资产的公平市场价格．第一种方法通过构建无风险投资组合得到衍生工具的价格．具体来说，我们用投资组合中对权重的无穷小调整和期权价格的变化复制标的资产 S_t 的不可预测变动．这种方法消除了投资组合的风险溢价，同时对 $F(S_t,t)$,S_t 的路径做出了限制，无风险资产可以随时间共同变化．这里可以对仓位做出无穷小调整的假设很重要，它说明了连续时间资产定价模型的优势． 261

衍生资产定价的第二种方法依赖于找到概率测度 Q，使得 $\mathrm{e}^{-rt}F(S_t,t)$ 在该测度下是鞅，即

$$\mathrm{e}^{-rT}F(S_t,t) = E^Q\left[\mathrm{e}^{-rT}F(S_T,T) \mid I_t\right], \quad t < T \tag{15.85}$$

换句话说，就是令随机微分

$$\mathrm{d}\left[\mathrm{e}^{-rt}F(S_t,t)\right], \quad 0 \leqslant t \tag{15.86}$$

的漂移项为零．

不管是直接求解 Black 和 Scholes 的基础 PDE，还是像上一节那样利用等价测度 Q 计算期望 $E^Q[e^{-rT}F(S_T,T)\,|\,I_t]$，我们都可以推导出 Black-Scholes 公式．在 Black 与 Scholes 的论文中使用的是第一种方法．上一节我们利用鞅方法得到了相同的结果，过程稍显冗长，但其中的数学运算却很简单．

显然，这两种方法有某种联系．我们在这一节中要讨论的就是这个问题．

这为应用前面介绍过的高阶数学工具提供了一个机会．特别地，我们讨论的内容也是如下条目的一个例子：

- 微分形式和积分形式伊藤引理的应用；
- 伊藤积分的鞅性质；
- Girsanov 定理的重要应用．

我们分两个步骤说明 PDE 方法与鞅方法之间的联系．第一步使用伊藤引理的符号形式，简洁、直观，但有很多数学上的问题没有确切的解答．这一步的重点在于 Girsanov 定理的应用．第二步我们则用到了积分形式的伊藤引理．

下面将针对 $e^{-rt}F(S_t,t)$应用伊藤引理，这要求 $F(\cdot)$关于 S_t 二阶可微，关于 t 一阶可微．这些假设在后面的讨论中不再重复．

15.5.1 两种方法的等价性

我们逐步说明两种方法之间的联系．我们首先说明如何通过变换维纳过程及相应概率测度将 $e^{-rt}S_t$ 变为鞅，然后对衍生资产价格 $e^{-rt}F(S_t,t)$进行相同讨论．

直接应用 Girsanov 定理就可以进行这些变换(当推导 Black-Scholes 公式时，将概率测度从 P 变到 Q 不需直接使用 Girsanov 定理)．

1. 将 $e^{-rt}S_t$ 变为鞅

我们先设定用来描述标的资产价格 S_t 动态变化的基本模型．在这里，我们假设标的资产价格服从随机微分方程

$$dS_t=\mu(S_t)dt+\sigma(S_t)dW_t,\quad t\in[0,\infty) \tag{15.87}$$

其中漂移项和扩散项仅与观察到的标的资产价格 S_t 有关．并假设这些系数满足通常的规范性条件，而 W_t 则是概率测度 P 下的维纳过程．

简化该随机微分方程使记法更清楚．将其写作

$$dS_t=\mu_t dt+\sigma_t dW_t \tag{15.88}$$

262 本章第一节直接找出了概率测度 Q，将 $e^{-rt}S_t$ 变为鞅．接下来我们利用 Girsanov 定理进行相同的操作．

计算按无风险利率贴现后的价格 $e^{-rt}S_t$，对 $e^{-rt}S_t$ 应用伊藤引理，有

$$d[e^{-rt}S_t]=S_t d[e^{-rt}]+e^{-rt}dS_t \tag{15.89}$$

展开 dS_t，我们可以得到

$$d[e^{-rt}S_t] = e^{-rt}[\mu_t - rS_t]dt + e^{-rt}\sigma_t dW \tag{15.90}$$

通常来说，由于 S_t 是风险资产[㊀]

$$[\mu_t - rS_t] > 0 \tag{15.91}$$

该等式漂移项不为 0，$e^{-rt}S_t$ 不是鞅．

但是我们可以利用 Girsanov 定理将 $e^{-rt}S_t$ 变为鞅．由于这是 Girsanov 定理在金融领域的基本应用．我们将详细地给出每一步计算．

Girsanov 定理表明，存在 I_t 适应的随机过程 X_t 和 W_t^*，使得

$$dW_t^* = dX_t + dW_t \tag{15.92}$$

且 W_t^* 在下面这个新的概率测度之下是一个维纳过程

$$dP = \xi_t dQ \tag{15.93}$$

其中 ξ_t 定义为

$$\xi_t = e^{\int_0^t X_u dW_u - \frac{1}{2}\int_0^t X_u^2 du} \tag{15.94}$$

假设随机过程 X_t 满足 Girsanov 定理中其他可积条件．[㊁]

为了达成我们的目的，(15.92)是一个很重要的等式．具体来说，我们可以用这个等式消去(15.90)中的 dW_t．代入 dW_t^* 后，该等式可以写成

$$d[e^{-rt}S_t] = e^{-rt}[\mu_t - rS_t]dt + e^{-rt}\sigma_t[dW_t^* - dX_t] \tag{15.95}$$

整理之后，可得

$$d[e^{-rt}S_t] = e^{-rt}[\mu_t - rS_t]dt - e^{-rt}\sigma_t dX_t + e^{-rt}\sigma_t dW_t^* \tag{15.96}$$

根据 Girsanov 定理，如果我们在新概率测度 Q 下考虑这个 SDE，W_t^* 会是标准维纳过程．此外，如果我们令漂移项等于 0，那么 Q 成为了鞅测度．为了做到这一点，我们只需要定义 dX_t 如下

$$dX_t = \left[\frac{\mu_t - rS_t}{\sigma_t}\right]dt \tag{15.97}$$

我们假设这样定义的 dX_t 满足 Girsanov 定理中要求的积分条件．

这样就得到了推导的第一步．我们得到了一个鞅测度 Q，及这个测度下一个新的维纳过程 W_t^*，还有一个相应的漂移调整 X_t 使得 $e^{-rt}S_t$ 变为鞅，且满足 SDE：

$$d[e^{-rt}S_t] = e^{-rt}\sigma_t dW_t^* \tag{15.98}$$

我们可以利用该式将 $e^{-rt}F(S_t,t)$ 变换为鞅．

2. 将 $e^{-rt}F(S_t,t)$ 变为鞅

前一节的推导得出了对衍生资产应用 Girsanov 定理所需的漂移调整 X_t 的确切形式．要对衍生资产定价，我们还需在本节中利用 Girsanov 定理证明 $e^{-rt}F(S_t,t)$ 在测度 Q 下也

㊀ 如果持有 S_t 美元无风险资产，那么在无穷小的时间 dt 内，将获得 $rS_t dt$ 美元收益，而 $\mu t dt$ 是对该资产的真实期望收益.

㊁ 即原始系统中漂移系数和扩散系数都是有定义的．

具有鞅性质．

证明的方法与之前的类似．首先，利用微分形式的伊藤引理得到关于 $e^{-rt}F(S_t,t)$ 的随机微分方程，然后对其中的维纳过程应用 Girsanov 变换．

直接求导，我们可以得到

263

$$d[e^{-rt}F(S_t,t)] = d[e^{-rt}]F + e^{-rt}dF \tag{15.99}$$

注意，我们将右边的 $F(S_t,t)$ 简写为 F. 使用伊藤引理，并代入 dF，得到关于 $d[e^{-rt}F(S_t,t)]$ 的 SDE：

$$d[e^{-rt}F(S_t,t)] = e^{-rt}[-rF\,dt] + e^{-rt}\left[F_t dt + F_s dS_t + \frac{1}{2}F_{ss}\sigma_t^2 dt\right] \tag{15.100}$$

现在的问题在于如何替代 dS_t. 有两种选择．由于 $e^{-rt}S_t$ 在 W_t^* 和 Q 测度下是鞅，可以使用

$$d[e^{-rt}S_t] = e^{-rt}\sigma_t dW_t^* \tag{15.101}$$

也可以使用(15.87)中的原始 SDE：

$$dS_t = \mu_t dt + \sigma_t dW_t \tag{15.102}$$

在这里我们选择第二种替换方式．这样的选择可以让我们再一次看到 Girsanov 定理具体该如何使用．现在，利用(15.102)消去(15.100)中 dS_t，

$$d[e^{-rt}F(S_t,t)] = e^{-rt}[-rF\,dt] + e^{-rt}\left[F_t dt + F_s[\mu_t dt + \sigma_t dW_t] + \frac{1}{2}F_{ss}\sigma_t^2 dt\right] \tag{15.103}$$

整理后可得

$$d[e^{-rt}F(S_t,t)] = e^{-rt}[-rF + F_t + F_s\mu_t + \frac{1}{2}F_{ss}\sigma_t^2]dt + e^{-rt}F_s\sigma_t dW_t \tag{15.104}$$

再次使用 Girsanov 定理．考虑维纳过程 W_t^*：

$$dW_t^* = dW_t + dX_t \tag{15.105}$$

我们随后可以对(15.104)中 SDE 进行变换：

$$\begin{aligned} d[e^{-rt}F(S_t,t)] =& e^{-rt}[-rF + F_t + F_s\mu_t + \frac{1}{2}F_{ss}\sigma_t^2]dt \\ & - e^{-rt}F_s\sigma_t dX_t + e^{-rt}F_s\sigma_t dW_t^* \end{aligned} \tag{15.106}$$

注意，(15.106)中误差项 dW_t^* 仅在概率测度 Q 下是标准维纳过程．因此，从现在开始，我们是在概率测度 Q 之下考虑这个 SDE.

dX_t 的值在(15.97)中已经给出：

$$dX_t = \frac{\mu_t - rS_t}{\sigma_t}dt \tag{15.107}$$

将它代入(15.106)：

$$\begin{aligned} d[e^{-rt}F(S_t,t)] =& e^{-rt}[-rF + F_t + F_s\mu_t + \frac{1}{2}F_{ss}\sigma_t^2 - F_s(\mu_t - rS_t)]dt \\ & + e^{-rt}F_s\sigma_t dW_t^* \end{aligned} \tag{15.108}$$

化简得

$$d[e^{-rt}F(S_t,t)] = e^{-rt}\left[-rF + F_t + F_s rS_t + \frac{1}{2}F_{ss}\sigma_t^2\right]dt + e^{-rt}F_s\sigma_t dW_t^* \quad (15.109)$$

但是要使 $e^{-rt}F(S_t,t)$ 在 W_t^* 和 Q 测度下是鞅，该 SDE 的漂移项必须为 0，[⊖] 因此我们必须有

$$-rF + F_t + F_s rS_t + \frac{1}{2}F_{ss}\sigma_t^2 = 0 \quad (15.110)$$

这个表达式与 Black 和 Scholes 给出的基础 PDE 相同．在此 dX_t 下，按无风险利率贴现的衍生品价格服从下面的 SDE：

$$d[e^{-rt}F(S_t,t)] = e^{-rt}\sigma_t F_s dW_t^* \quad (15.111)$$

漂移参数为 0.

15.5.2 推导的关键步骤

264

在该推导过程中，有几个关键步骤需要进一步讨论．

首先是 Girsanov 定理的使用方法．对于按无风险利率贴现的金融资产价格，已知一个维纳过程驱动的 SDE，价格过程在开始时不是鞅，我们的目标是将它变为鞅．

我们利用 Girsanov 定理找到新维纳过程和新概率测度 Q，使贴现后的资产价格变为鞅．概率测度 Q 被称为等价鞅测度．我们还可以得到 Girsanov 定理需要的漂移调整项 X_t．在前面的推导过程中，我们在(15.96)和(15.106)中使用了两次 Girsanov 定理．

这将带到出了推导过程中第二个关键点．回顾(15.106)：

$$\begin{aligned} d[e^{-rt}F(S_t,t)] =& e^{-rt}[-rF + F_t + F_s\mu_t + \frac{1}{2}F_{ss}\sigma_t^2]dt \\ & - e^{-rt}F_s\sigma_t dX_t + e^{-rt}F_s\sigma_t dW_t^* \end{aligned} \quad (15.112)$$

代入 dX_t 的值，即在漂移项中加上

$$dX_t = \frac{\mu_t - rS_t}{\sigma_t} \quad (15.113)$$

注意该变换起到的微妙作用．该 dX_t 使(15.104)的漂移项中的 $F_s\mu_t dt$ 被消去，变成 $F_s r dt$.

也就是说，应用 Girsanov 定理相当于将漂移项 μt 变为 rS_t. 衍生品方面的论著通常会将所有漂移系数都替换为无风险利率 r. 这里我们可以清楚地看出，Girsanov 定理为这种变换提供了理论基础．

最后一点也是第三点，如何找到可以同时将 $e^{-rt}S_t$ 和 $e^{-rt}F(S_t,t)$ 变为鞅的 W_t^* 和 Q 呢？这个问题很重要，因为鞅的函数不一定还是鞅．

这一步与金融资产的均衡定价和套利定价有关，属于动态资产定价理论的范畴．我们只对基本原理进行简要提及．第 2 章讨论过，在合适条件下，资产价格之间的套利关系可以产生唯一的鞅测度，使得所有(按无风险利率贴现后的)风险资产价格在此测度下都为鞅.

[⊖] 如果套利机会不存在，那么相同的 Q 将所有资产价格都变为鞅．

因此，在 Girsanov 变换中使用同一对W_t^*和 Q 是资产定价理论的必然结果．如果市场中存在套利机会，那么我们就不能这样做．

15.5.3 伊藤公式的积分形式

前面通过微分形式的伊藤引理对 PDE 方法与鞅方法之间的联系进行了讨论．值得注意的是，这个微分形式的伊藤引理只是一个符号化的表达．

在之前的章节，我们好几次强调过，我们考虑的随机微分属于符号形式，其背后真正表达的是一个积分等式．而所有 SDE 其实都是建立在这种伊藤积分的基础上的．我们使用随机微分是为了更方便地表达冗长的方程．

利用积分形式的伊藤引理也可以进行相同的分析．不回顾全部细节，我们只重复基本的步骤.

按无风险利率贴现的看涨期权价格仍用 $e^{-rt}F(S_t,t)$表示．对它应用积分形式的伊藤引理：

$$e^{-rt}F(S_t,t)=F(S_0,0)+\int_0^t e^{-ru}\left[-rF+F_t+\frac{1}{2}F_{ss}\sigma_u^2+F_s rS_u\right]du+\int_0^t e^{-ru}\sigma_u F_s dW_t^* \tag{15.114}$$

265 注意，我们使用的是 W_t^* 而不是 W_t，并用无风险利率 r 替代了 μ_t.

假设 σ_t 满足

$$E^Q\left[e^{\int_0^t (F_s e^{-ru}\sigma_u)^2 du}\right]<\infty \tag{15.115}$$

这是 Girsanov 定理中的 Novikov 条件，说明积分

$$\int_0^t F_s e^{-ru}\sigma_u du \tag{15.116}$$

在测度 Q 下是鞅．

但是按 e^{-rt} 贴现的衍生资产价格也是鞅．这使(15.114)右边第一个积分项

$$\int_0^t e^{-ru}\left[-rF+F_t+\frac{1}{2}F_{ss}\sigma_u^2+F_s rS_u\right]du \tag{15.117}$$

也是(平凡)鞅．但这个积分是关于时间的积分，而鞅不应有非零漂移系数．因此，该积分必须为 0，这就给出了偏微分方程

$$-rF+F_t+\frac{1}{2}F_{ss}\sigma_u^2+F_s rS_u=0,0\leqslant S_t,\quad 0\leqslant t\leqslant T \tag{15.118}$$

再次得到 Black 和 Scholes 的基本 PDE.

15.6 小结

本章讨论 Girsanov 定理的应用．我们对其中几个技术要点进行了讨论，主要结论如下.

在对衍生资产定价时，在鞅方法和 PDE 方法之间存在等价性．

如果使用鞅方法，我们在等价鞅测度下计算条件期望，而这种测度可以将所有按无风险利率贴现的资产价格变为鞅．只要理解了 Girsanov 定理的深层含义，就很容易对这些条件期望进行形象化理解．对于欧式衍生资产，这些期望也为寻找无套利价格的数值解提供了简单方法．

可以证明鞅方法与 PDE 法都能得到同样的 PDE. 差别在于：鞅方法中 PDE 是风险中性资产定价的结果，而在 PDE 方法中，我们从这些 PDE 出发计算无风险价格．

15.7 参考阅读

本节中我们推导 Black-Scholes 公式的部分参考了 Ross(1993)的方法．Cox 和 Huang (1989)对主要鞅结果有着精彩的综述，方法与 Duffie(1996)相同．

15.8 习题

1. 我们在本题中用 Girsanov 定理对选择期权进行定价．选择期权是一种奇异期权，它赋予持有者在未来某时刻选择看涨期权还是看跌期权的权利，标的资产相同．T 表示到期日，S_t 表示股票价格，K 表示执行价格．如果我们在 t 时刻买入选择期权，此时标的资产价格为 S_t，我们可以选择执行价格为 K、标的价格为 S_t 的看涨期权或看跌期权．t 时刻看涨期权的价值为

$$C(S_t,t) = e^{-r(T-t)}E[\max(S_T-K,0)\mid I_t]$$

而看跌期权的价值为

$$P(S_t,t) = e^{-r(T-t)}E[\max(K-S_T,0)\mid I_t]$$

且 t 时刻时，选择期权的价值为：

$$H(S_t,t) = \max[C(S_t,t),P(S_t,t)]$$

(a) 利用上述条件，证明：

$$C(S_t,t)-P(S_t,t) = S_t - Ke^{-r(t-t)}$$

266

这使你想到著名的平价条件了吗？

(b) 下面证明 t 时刻选择期权的价值为

$$H(S_t,t) = \max[C(S_t,t),C(S_t,t)-S_t+Ke^{-r(T-t)}]$$

(c) 最后，证明 0 时刻期权价值为

$$H(S_0,0) = C(S_0,0)+e^{-rT}E[\max[K-S_0e^{rT}e^{\sigma W_t-\frac{1}{2}\sigma^2 t},0]]$$

其中 S_0 为 0 时刻观察到的标的价格．

(d) 现在来看使用 Girsanov 定理的地方．如何使用 Girsanov 定理，并轻松地计算上述期望的值？

(e) 写出选择期权的最终公式．

2. 我们在本题中将 Black-Scholes 条件应用到外汇资产中，可以看到 Girsanov 定理的不同应用[更具体的细节见 Musiela 和 Rutkowski，1997]．令 r，f 分别表示国内外的无风险利率．S_t 表示汇率，即用本币表示的每单位外币价格．假设 S_t 服从几何布朗运动：

$$dS_t = (r-f)S_t dt + \sigma S_t dW_t$$

(a) 证明

$$S_t = S_0 e^{(r-f-\frac{1}{2}\sigma^2)t+\sigma W_t}$$

其中 W_t 是在概率测度 P 下的维纳过程．

(b) 随机过程

$$\frac{S_t e^{ft}}{S_0 e^{rt}} = e^{\sigma W_t - \frac{1}{2}\sigma^2 t}$$

是测度 P 下的鞅吗？

(c) 令 Q 表示概率测度

$$Q(A) = \int_A e^{\sigma W_T - \frac{1}{2}\sigma^2 T} dP$$

在 Q 测度下，对于随机过程 $W_t - \sigma t$ 来说，Girsanov 定理意味着什么？

(d) 利用伊藤公式证明

$$dZ_t = Z_t[(f - r + \sigma^2)dt - \sigma dW_t]$$

其中 $Z_t = 1/S_t$.

(e) 随机过程 $Z_t e^{rt}/e^{ft}$ 在哪种测度下是鞅？

(f) 我们可以认为 Q 测度在外国市场上也是一个无套利测度吗？

3. 考虑下面的 SDE，其中 S_t 为实际测度下的过程：$dS_t = \mu dt + \sigma dW_t$. 写出风险中性测度下相应动态过程，使其贴现后的股票价格为鞅．在风险中性测度下，使用本书中对数模型推导的方法来得出看涨期权的闭形值．

267 **4.** 写出计算欧式数字期权的 Matlab 程序，不要使用拟合法．

第 16 章　利率敏感型证券的新结论和工具

16.1　引言

本书的第一部分介绍了经典 Black-Scholes 方法中实用的量化工具．我们假设标的证券 S_t 是无股利支付的股票，无风险利率 r 和隐含波动率 σ 都是常数，期权为欧式，且不考虑交易费用和不可分割性．

金融市场上交易的衍生品远比普通的看涨看跌期权要复杂，而我们介绍的简化的模型只适用于那些简单的衍生品．事实上，Black-Scholes 使用的一些假设，尽管具有较好的鲁棒性，但可能对利率敏感证券并不适用．[㊀] 为此，我们需要一些新的假设并且建立更复杂的模型．

这类新的衍生品在很多方面与我们已经介绍的普通衍生品有相似之处，但是也有一些独特的难点．更重要的是，利率敏感工具和利率期限结构的研究给我们带来了一些新的结果．当然，为了得到这些结果，我们需要一些不同的量化工具．

回顾前几章，可以发现我们介绍的例子很大程度上与基本 Black-Scholes 框架有关，特别是以下两个方面：

1）我们没有讨论可提前执行的美式衍生证券．

2）我们总是假设无风险利率 r 恒等于一个常数．

然而，对于大部分金融衍生品而言，以上两点都是过强的限制．[㊁]

首先，大部分金融衍生品都是美式的，可以提前执行．金融衍生品的持有者通常不一定等到到期日才执行期权．这显然使得定价的难度陡增，我们必须为此引进新的数学工具．

第二，现实金融市场中，无风险利率显然并不是常数．和其他金融资产的价格一样，利率也具有不可预测、无穷小的波动．对于一些金融衍生品如股票期权，假设利率为常数并不完全正确，但仍然是一种合理近似．然而，对于利率衍生品来说，我们就不能采用这个假设了．正是利率变化中不可预测的风险使得这类衍生品在市场上受到欢迎．因此，我们需要对无风险利率模型引入不可预测的维纳成分，而这将会使所需的数学工具更加复杂．

最后注意到，当衍生品期限较短时，我们可以使用 Black-Scholes 假设．而当期限较长时，我们就需要放宽对利率和波动率为常数的假设．

本书的第二部分将介绍进行这些调整所需的新工具以及建立结构模型所需要的重要新结果．

㊀ Black-Scholes 假设的鲁棒性是该公式至今仍受到市场专业人士欢迎的原因．例如，当实际中波动率为随机，或者在利率随机移动的情况下，我们仍能够得到相对准确的价格．具体的研究可以参见 El Karoui 等(1998)．

㊁ Merton(1973)首先尝试引入随机利率．不过标的资产仍然是股票．这种调整所得的问题仍然能用经典工具加以解决．而当衍生品关于利率非常敏感时，新工具的实用性才得以体现出来，此时支付情况依赖于利率的值或轨迹．

16.2 概要

本章我们简要介绍一下这些新工具背后的基本思想．接下来几章介绍的内容更为高阶，它们对高流动性的衍生品定价更为实用．

第 17 章将重新讨论本书第一部分中的简单双状态模型．不同之处在于，我们将采用更复杂的初始设定来推导出一个新的模型．具体来说，在这个新模型中，我们假设的短期利率是随机的，而研究的对象是对利率变化敏感的衍生品．这个新模型会是第 2 章中那个简单模型的新版本．通过这些讨论，我们能够研究标准化等重要概念和远期测度等工具．

第 18 章则是为利率期限结构的建模打好基础．我们将正式给出远期利率、即期利率、期限结构的定义．更重要的是，这一章将介绍两种对期限结构建模的主要方法：经典方法和 Health-Jarrow-Mor ton 方法．通过学习这两种方法在假设、内涵思想、实际应用上的区别，我们能够将它们应用到各种不同的情况．这是学习利率敏感型工具定价的重要的一步．

第 19 章介绍了利率敏感型证券的经典 PDE(偏微分方程)分析．这种方法可以视为对 Black-Scholes PDE 进行模仿，进而得到适用于无违约风险零息债券及其衍生品价格的 PDE. 在这个过程中，最困难的问题是如何调整短期利率过程中的漂移项．短期利率不是一项资产，因此它的漂移项不能像 Black-Scholes 模型那样直接用无风险即期利率 r 代替，而需要采取一种更复杂的处理方法．为此，我们需要引入利率风险的市场价格这一概念．最后我们得到的 PDE 能够解释这一新增的(还未观察到的)变量．

270

第 20 章将介绍固定收益类产品定价的经典 PDE 方法．第 21 章将介绍对利率敏感证券进行定价、对冲、套利的前沿工具．首先我们会讨论一类随机过程的条件期望与偏微分方程的基本关系．了解这种关联后，金融市场参与者能够得到一种非常重要且实用的工具．这种工具本身与 Feyman-Kac 公式有关，我们会在本章里仔细讨论它．分析师们在面对实际问题时，有两种可能的处理方法：不管是计算等价鞅测度下的条件期望，还是直接去求解对应的 PDE. 而这种关联性让分析师可以灵活地选择其中一种计算复杂度更低的方法来进行数值运算．

第 21 章还将介绍随机过程生成器的概念、Kolmogorov 后向方程，以及 Markov 性质的应用．后者对于短期利率模型特别重要．因为我们会发现短期利率模型并不是 Markov 过程，而这一点会使得对 Feynman-Kac 公式的引用变得更复杂．

最后，第 22 章将介绍停时的概念，这是美式衍生品的重要概念．介绍这一概念的同时我们将介绍一个非常重要的最优化算法——动态规划算法．本章我们还将展示美式证券二叉树模型与停时之间的关联．我们将看到，美式衍生品的定价是以动态规划的应用为基础的.

停时也是随机变量，它的取值表示某个随机过程终止的时间点．例如，美式看涨期权可以在到期日之前提前执行．最初时刻，执行时间是未知的．所以，期权的执行日期可以看作一个随机变量．停时为这种随机变量对定价的影响提供了数学工具．

这些数学工具在利率敏感型衍生品定价中特别实用．所以，在对这些工具进行详解讨

论前，我们先简单地了解一下它们．

16.3 利率衍生品

通过改变 Black-Scholes 的基本假设可以得到一系列更复杂的衍生工具，其中重要的一类就是以利率敏感型资产为标的的衍生品．

以下是一些著名的利率衍生品：㊀

- **利率期货与远期合约**．令 L_{t_i} 表示某贷款从 t_i 时刻到 t_{i+1} 时刻的年化单利率．假设不存在买卖价差和违约风险．那么，在 t 时刻，$t_i<t<t_{i+1}$，我们可以写出建立在"Libor 利率"㊁ L_{t_i} 上的期货与远期合约．例如，某$[t_i, t_{i+1}]$期间的贷款合同在 t 时刻签订，合同标定利率为 F_t．该远期合约的购买者将在 t_i 时刻得到一笔总值为 N 的贷款，并在 t_{i+1} 时刻偿还 $N(1+F\delta)$，其中 δ 是根据天数得出的调整因子．㊂

271

- **远期利率协议(FRA)**．早在第 1 章我们就已介绍过，这一类合约能够更方便地对冲利率风险．FRA 的持有者将在 t_{i+1} 时刻得到或支付差额．如果 $F_t>L_{t_i}$，他将得到 $N[F_t-L_{t_i}]\delta$；如果 $F_t<L_{t_i}$，他将支付 $N[L_{t_i}-F]\delta$．在 t 时刻签订 FRA 合约时选定合适的 F_t，从而使得该合约在 t 时刻价值为 0．图 16-1 描述了这种情况．在现实金融市场中，FRA 经常在 L_{t_i} 被观察到的同时，即 t_i 时刻就进行差额支付．因此，支付额需要用$(1+L_{t_i}\delta)$折现．图 16-1 也表明了这一点．

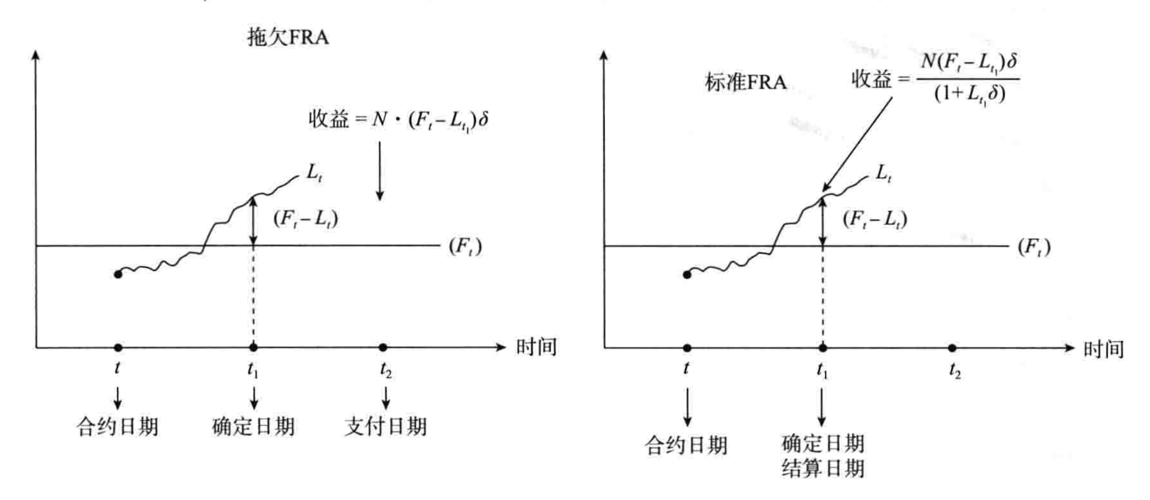


图 16-1 远期利率合约

㊀ 接下来，读者会注意到记号发生了细微的改变．具体而言，时间下标将用 t_i 表示．在新工具中我们也将采用这种记号．

㊁ Libor 是伦敦同业银行拆借利率．它是由资金卖家提供的银行内部利率．它是选择伦敦部分银行并计算它们给出的利率的平均值．所以，根据所选银行的不同，可能会有若干个相同期限的 Libor．英国银行家协会计算出"官方的"Libor，这也是大多数 Libor 衍生工具的基础．

㊂ 例如，它等于$[t_i, t_{i+1}]$期间的天数除以 365 所得的值．

- **上限和下限协议**．大多数流动性利率衍生品中都含有上限和下限协议．上限协议可以用来对冲利率升高的风险．而下限协议对冲利率下降的风险．它们实际上是建立在 Libor 利率上的篮子期权．假设 t 表示当前时刻，t_0 为上限协议的开始日期，$t\leqslant t_0$；t_n 为该协议的结束日期，n 为某个固定的数．令 $t_1, t_2, \cdots, t_{n-1}$ 为重置日．将上限协议应用到每个时间段$[t_i, t_{i+1}]$，协议的买方将在 t_{i+1}时刻收到

$$N \max[\delta(L_{t_i} - R_{\text{cap}}), 0] \tag{16.1}$$

其中 L_{t_i} 为 t_i 时刻观测到的 Libor 利率，δ 为根据天数得到的时间长度，N 为 t 时刻合约内的名义金额．R_{cap} 为上限利率，其作用与执行价格类似．也就是说，在一个时间段内，合约买方可以将利率锁定在 R_{cap} 以下．所以，它等价于一个到期日为 t_i、执行价格与 R_{cap}有关的看跌期权，而该期权标的物是到期日为 t_{i+1}、无违约风险的贴现债券．特别地，该期权的行权价格为 $100/(1+R_{\text{cap}}\delta)$．其中 R_{cap} 是上限利率，δ 为根据天数得到的时间长度．

272

该公式已由图 16-2 给出．在 t_i 时期权到期．如果债券价格低于 $100/(1+R_{\text{cap}}\delta)$，期权持有方将收到价格差：

$$\text{收益} = \frac{100}{(1+R_{\text{cap}}\delta)} - \frac{100}{(1+L_{t_i}\delta)}$$

否则，期权持有者什么也得不到．该过程也等价于一份关于 L_{t_i}、执行价格为 R_{cap} 的看涨期权，如图 16-2．这里，如果将每个上限协议视为到期日为 t_{i+1} 的看涨期权并建立在 Libor 利率 L_{t_i} 上，那么我们需要记住所有结算在 t_{i+1}时刻完成，而非 t_i 时刻．同样，利率下限协议可以视为到期日为 t_i 的看涨期权，而标的物是到期日为 t_{i+1} 的折现债券．与之等价，它也可以被看作到期日为 t_{i+1}，建立在 Libor 利率上的看跌期权．

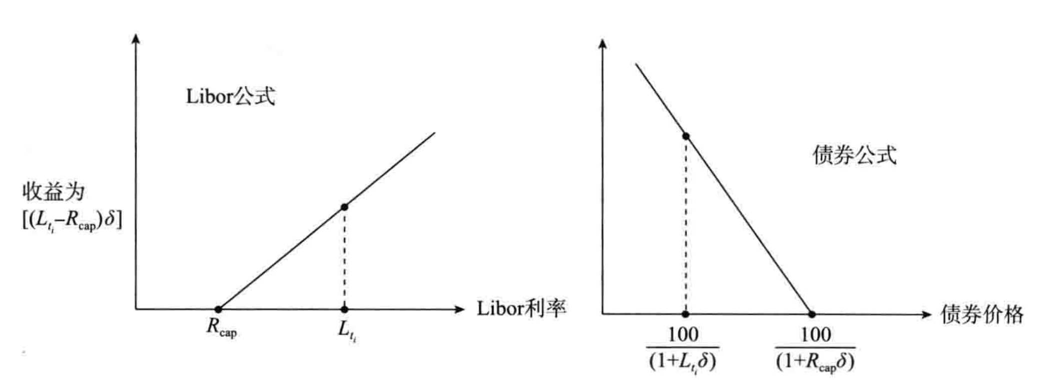


图 16-2 利率上限协议

- **利率互换**．我们在第 1 章也提到过这类工具．普通利率互换一般指双方交换一笔借款，本金相同，一方提供固定的事先设定的互换利率 κ，另一方提供的则是浮动的 Libor 利率 L_{t_i}．本金总量为 N，双方在 t_{i+1}时刻支付利息的差额．显然，利率互换是一种 FRA 的组合．我们需要选择合适的互换利率 κ，使得互换合约在 t 时刻的

价值为 0.

- **债券期权**. 如果一方持有以债券为标的物的期权，那么他可以以执行价格 K 买入价值为 B_t 的债券. 由于债券价格依赖于即期利率和远期利率，债券期权对 r_t 和 L_{t_i} 的变化都非常敏感.
- **互换期权**. 互换期权是建立在互换合约上的期权. 由于涉及多个到期日，这种期权很灵活. 例如，在 t 时刻，参与者买入此类期权，执行价格为 κ，名义金额为 N. 该期权在 T 时刻到期，而互换交易从 T_1 时刻开始，$T\leqslant T_1$，到 T_2 时刻结束，$T_1<T_2$. 互换期权的购买者在期权到期日可以参与互换交易，互换交易的利率为 κ、名义金额为 N，从 T_1 到 T_2. 所以，如果在 T_1 时刻互换利率 R_{t_i} 高于 κ，那么此时互换期权的价值大于零.

以上都是基础的利率衍生品. 本书篇幅有限，无法介绍其他一些更复杂的衍生品. 接下来我们将总结这些工具的关键要点，并思考需要哪些新的数学工具.

16.4 难点

273

利率衍生品为我们带来了几个亟待解决的问题，债券期权可以很好地体现出来. 然后我们将它与 Black-Scholes 框架比较.

债券 B_t 的价格依赖于市场上当前和未来即期利率的随机行为. 所以，我们需要两个新的假设：(1)债券价格 B_t 必须是当前和未来即期利率的函数.(2)即期利率 r_t 不能假设为常数，否则意味着 B_t 完全可预测、标的证券的波动率为 0，这样就不再需要以此为基础的看涨或看跌期权.

所以，首先我们需要引入随机利率. 但是，由于折现因子和相应的支付都与利率有关，这将使得无套利定价更加复杂.

第二个难点是大多数利率衍生品是美式的，允许在到期日前被执行.

第三，对于利率衍生品而言，标的证券所产生的现金流与一般利率衍生品不同. 举个例子，如果市场上存在盯市制度(marking-to-market)，即期利率的随机性通常会造成期货合约的无套利定价不同于远期合约的无套利定价. 这是盯市制度所要求的定期结算造成的. 随着利率波动，盯市制度规范下的期货合约持有者会定期地获得或者支出某个数量的现金流. 而当我们对这些由于盯市而产生的现金流做贴现处理时，用的也是随机利率. 这就会导致期货价格与不受盯市规范的远期合约价格相比有所不同.

类似地，如果债券会支付息票，标的证券 B_t 也会与不支付股利的股票 S_t 有所区别.

以上是在研究利率衍生品时需要做出的重要调整. 此外还有一些表面上不明显的技术细节，下面介绍其中一点.

16.4.1 漂移项调整

利率不是资产，它们更像是资产的“回报”. 原先我们对于资产价格过程的处理方式，即用无风险利率 r 来代替未知的漂移项来满足无套利条件的做法在这里就不可行了. 实际

上，r_t 的动态变化可用 SDE(随机微分方程)表示为：

$$\mathrm{d}r_t = a(r_t, t)\mathrm{d}t + \sigma(r_t, t)\mathrm{d}W_t$$

而其中的漂移系数 $a(r_t, t)$ 必须用其他方法进行风险调整．这使得 Girsanov 定理的使用不再像之前那么直接．事实上，在风险中性测度 Q 下，我们不能直接把 W_t 变换成新的维纳过程 W_t^*．我们来看一下为什么：已知两个维纳过程的 Girsanov 存在对应关系：

$$\mathrm{d}W_t = \mathrm{d}W_t^* - \lambda_t \mathrm{d}t$$

将其替换入原先的 SDE 中，可得：

$$\mathrm{d}r_t = (a(r_t, t) - \lambda_t\sigma(r_t, t)\mathrm{d}t) + \sigma(r_t, t)\mathrm{d}W_t^*$$

而现阶段，我们并不清楚 λ_t 的取值是什么．

这比 Black-Scholes 模型中的内容的替换明显更复杂了：当 S_t 为可交易资产的价格时，我们只需要做如下的替换

$$r = (a\mu - \lambda_t\sigma) \tag{16.2}$$

274 金融基本理论可以说明这一点．由于利率 r_t 不是一项“资产”，类似的替换可能不成立．

16.4.2 期限结构

利率衍生品定价的另一个困难之处在于市场上同时存在着种类繁多的利率．在简单的 Black-Scholes 体系中，只考虑单个标的资产 S_t．但在固定收益产品市场上，根据期限的不同，存在着各种不同的利率．而且这些不同的利率的动态变化规律差别不大，更重要的是，它们彼此之间存在很强的相关性．

因此，与 Black-Scholes 模型中的利率相比，我们要面对的是一个更复杂的对象：一个随机过程构成的向量．这个向量的各个分量之间就必须满足复杂的关系，以消除套利的可能性．所得的 k 维动态系统因而会更加复杂．

在经典 Black-Scholes 环境下，对标的资产价格过程的建模只含有一个 SDE，我们只需要考虑对单个变量在全过程的套利进行限制．但对于利率，同样的限制施加到 k 个变量．这不只是仅仅增加了变量数，我们还需要考虑若干不同利率组合的套利可能性．

最后一点是对波动率的建模．债券的波动率必须随时间不断变化．债券会在某个时刻到期，这一点和股票不一样．因此它的波动率不可以设为常数．

显然，利率衍生品不能用 Black-Scholes 环境下的假设来处理．

在处理利率衍生品时，刚刚提到的一部分困难可以通过小幅修改 Black-Scholes 原有假设并使用一些技巧就能克服．但是提早执行的可能性以及随机利率这两项改变所带来的困难，必须要引进一些新的数学工具才能解决．接下来的章节将解决这些问题．

16.5 小结

本章只是简要的概括，不能视为利率衍生品的正式介绍．不过它能够帮助我们理解接下来几章所介绍的工具．

16.6 参考阅读

Risk 所著的《Vasicek and Beyond》是利率衍生品及其定价方面我们极力推荐．读者还可以参考 Hull(2009)或者 Rebonato(1998)，后者含有更深入的研究内容．

16.7 习题

1. 画出以下金融工具的收益图：

(a) 即将到期的利率上限协议，上限利率为 $R_{cap}=6.75\%$，建立在 3 个月后的 Libor 利率 L_t 上．

(b) 远期合约，标的物为无违约风险的折现债券，债券在 2 年后到期，合约本身是在 3 个月后到期，合约价格为 89.5.

(c) 一个 3×6 的 FRA，协议固定利率(期限为 3 个月)为 F，而参考利率是 3 个月后的 3 个月期 Libor.

(d) 利率互换，支付固定利率方互换利率为 $\kappa=7.5\%$. 合约的期限为 2 年，作为互换的浮动利率为 6 个月期 Libor. 合约的开始日期为 6 个月前．

(e) 互换期权，期权持有方可以在 6 个月后成为一个 2 年期的利率互换的固定利率支付方，其协议固定利率为 $\kappa=0.6\%$

2. 以下哪种资产可以在金融市场交易：

(a) 6 个月期的 Libor

(b) 5 年期的国库券

(c) FRA 合约

(d) 上限协议

(e) 30 年期德国国债的收益

(f) 联邦基金利率的波动率

(g) 利率互换

3. 说明选择均值回归过程对利率进行建模的理由．为什么这种方法对利率合适，而对股票不合适？最基本的区别在哪里？

4. 考虑风险中性测度下的 Vasicek 模型：

$$dr_t = \alpha(\mu - r_t)dt + \sigma dW_t \tag{16.2}$$

其中 $r_0=0.01$，$\alpha=0.2$，$\mu=0.01$，$\sigma=0.05$.

编写 Matlab 程序，模拟一条 r_t 的样本路径，从 $t=0$ 到 $t=1$.

275 ~ 276

第 17 章　新环境下的套利定理

17.1　引言

在第 2 章中我们通过简单模型，介绍了在衍生品定价中使用的主要工具的缘由．当时我们介绍了综合(鞅)概率，它在本书的第一部分起着核心的作用．由于当时的假设很简单，它与一些复杂概念如风险中性概率、鞅工具都非常契合．

在第 2 章的模型中，以常利率进行无风险借贷是 3 种投资手段之一，另外两种分别是股票和建立在股票上的期权．在整个第 2 章，利率被假设为常数，我们没有讨论利率敏感型的金融衍生品．但是，在金融市场绝大多数交易品都是利率敏感的．它们常用来对利率风险进行对冲、套利或投机，因此，放宽对常数利率的假设是极为必要的．正如第 16 章所说，放宽常利率假设并引入复杂利率衍生品之后，我们需要新的数学工具，其中大多数是较为前沿的研究成果．

本章将沿这一方向继续推进．我们通过一个与第 2 章类似的离散模型，介绍新的数学工具. 我们将该模型从一个新的方向上加以拓展，同时确保这些新的工具和概念易于理解. 通过拓展第 2 章的简化模型，我们至少能够得到 3 个重要的新结论．

首先我们需要进行一些标准化处理．这可以用来得到资产价格比率的方程，但不是资产价格的方程．比率包含分子和分母．在动态环境下，这些都会发生变化．每个元素的期望变化率可能是未知的，但在一定条件下，两个元素之比的期望变化率可能是已知的．而比率会保持不变．所以，如果仔细选择定价公式中的分子和分母，我们就能有技巧地使用 Girsanov 定理，从而大大简化资产定价的建模．

在介绍标准化之前，我们首先假设短期利率在从一个时间段到另一时间段时会发生随机波动，讨论在这样的设定下，第 2 章中的基本结论是否还成立．显然，这个假设使得模型符合利率衍生品的特点，因为对这种证券定价需要假设利率是随机的．但这还不是最主要的．

结果显示一旦利率变为随机，我们就有新的方法寻找综合概率，特别是当我们研究利率敏感型工具时．虽然第 2 章中的主要思想没有发生太大变化，操作方法却变得截然不同. 事实上，对于不同的金融衍生品，采用不同的综合概率，可能会使得方法变得更具可操作性．而在各个情况下最终得到的无套利价格都是一样的．毕竟，起决定性作用的不是综合概率，而是其独有的状态价格向量．只不过某些综合概率可能比其他更加便于计算．

这一步骤很简单，但它绝对不像表面上看上去的那样无关紧要，特别是对于在金融中把综合概率测度作为定价工具的应用而言．实际上我们会发现，在定价过程中选择某个测度可能会远比选择另一个同样正确的测度更加简单，起到事半功倍的效果．本章的第二个

目标即是在简单的条件下解释这种复杂的思想.㊀

此外，之前几章内容所能涵盖的衍生品数量非常有限. 大多数讨论都是围绕 Black-Scholes 环境下的普通期权，只有少部分介绍了远期合约. 本章意味着我们的研究上了一个新的阶梯. 建立在 Libor 利率和债券上的衍生品是流动性最强的，但这些不能用第 2 章中即期利率为常数的模型来研究. 本章我们将通过金融基本定理来介绍这些重要金融工具，并说明引入新数学方法的必要性.

17.2 新金融工具的模型

我们先回忆第 2 章中的简化设定. 在单时间段、双状态的设定下，我们考虑无股利支付的股票 S_t、欧式看涨期权 C_t 以及无风险借贷款. 金融基本定理给出了这 3 种资产未来价值和当前无套利价格的线性关系： 278

$$\begin{bmatrix} 1 \\ S_t \\ C_t \end{bmatrix} = \begin{bmatrix} (1+r\Delta) & (1+r\Delta) \\ S_{t+\Delta}^u & S_{t+\Delta}^d \\ C_{t+\Delta}^u & C_{d+\Delta}^d \end{bmatrix} \begin{bmatrix} \psi^u \\ \psi^d \end{bmatrix} \tag{17.1}$$

其中 Δ 为两个时间点之间的间隔，u 和 d 代表两个状态，$\{\psi^u>0,\psi^d>0\}$ 为状态价格. 第一行表示无风险借贷的报酬，第二行表示股票 S_t，第三行表示期权 C_t.㊁

根据金融基本定理，对于已知的 $\{r_t,S_t,C_t\}$，在无套利的假设下，存在大于零的 $\{\psi^u,\psi^d\}$. 反之亦然. 如果 $\{\psi^u,\psi^d\}$ 存在且大于零，那么不存在套利机会.

无风险概率 Q 可以通过矩阵的第一行得到：

$$1=(1+r\Delta)\psi^u+(1+r\Delta)\psi^d$$

定义

$$Q_u=(1+r\Delta)\psi^u$$

可得

$$Q_d=(1+r\Delta)\psi^d$$

已知存在大于 0 的状态价格 ψ^u、ψ^d，$0\leqslant Q_u$，$0\leqslant Q_d$ 成立.

所以 Q_u 和 Q_d 是两个和为 1 的正数. 这满足简单环境下概率分布的要求，我们称 Q_u 和 Q_d 是综合概率，或者更具体地说是风险中性概率. 这种概率没有告诉我们关于真实世界中状态 u、d 的几率，它们是“风险中性”的，原因如下：

将系统(17.1)的第二和第三行单独提取出来：

$$S_t=S_{t+\Delta}^u\psi^u+S_{t+\Delta}^d\psi^d \tag{17.2}$$

$$C_t=C_{t+\Delta}^u\psi^u+C_{t+\Delta}^d\psi^d \tag{17.3}$$

用 $(1+r\Delta)/(1+r\Delta)$ 乘以 ψ^u 和 ψ^d，利用 Q_u 和 Q_d 得到以下定价方程：

㊀ 我们还可以提出以下问题：既然可以修改真实概率分布，从而使资产价格成为鞅，什么才是最好的修改方式呢？

㊁ 相比第 2 章，我们在记号上作了一点小的修改. 我们引入了上标 μ 和 d，代表两种状态.

$$S_t = S_{t+\Delta}^u \frac{Q_u}{1+r\Delta} + S_{t+\Delta}^d \frac{Q_d}{1+r\Delta} = \frac{1}{1+r\Delta} E^Q[S_{t+\Delta}] \tag{17.4}$$

以及

$$C_t = C_{t+\Delta}^u \frac{Q_u}{1+r\Delta} + C_{t+\Delta}^d \frac{Q_u}{1+r\Delta} = \frac{1}{1+r\Delta} E^Q[C_{t+\Delta}] \tag{17.5}$$

其中 $E^Q[\cdot]$是使用 Q_u、Q_d 作为概率的条件期望算子．为方便起见，本章忽略了 $E^Q[\cdot]$ 的下标 t.

根据这些定价方程，风险资产的期望回报用无风险利率折现之后，得到了无套利现价. 在这种意义下，Q_u、Q_d 是风险中性的．虽然 S_t、C_t 包含风险溢价，但这并不妨碍它们使用 Q_u、Q_d，就好像它们是来自风险中性世界一样．

从这些定价方程中还能得到另一个重要结果．重新组合(17.4)和(17.5)，可得

$$1 + r\Delta = E^Q\left[\frac{S_{t+\Delta}}{S_t}\right]$$

279

$$1 + r\Delta = E^Q\left[\frac{C_{t+\Delta}}{C_t}\right]$$

概率 Q 修正了风险资产的期望收益，使得所有期望收益率等于无风险利率 r. 所以我们将 Q 称为无风险概率或测度．

下一节我们将该框架从两个方面扩展．首先，我们将增加一个时间段，这样就可以引进即期利率随机波动的影响．第二，我们会改变所研究金融工具的类型，引入利率敏感型证券.

17.2.1 新环境

考虑两个时间段，记为 $t_1 < t_2 < t_3$，在每个单独的时间段中保留双状态的假设．增加时间段能够同时增加结果的数量．在 $t_1 = 1$ 时刻考虑 $t_3 = 1 + 2\Delta$ 时刻的情况，将有 4 个可能的状态，$\{\omega_i, i = 1, \cdots, 4\}$，表示价格在时间节点$\{t_1, t_2, t_3\}$之间的所有可能路径：

$$\{\omega_1 = \text{下降,下降}; \omega_2 = \text{上升,下降}; \omega_3 = \text{下降,上升}; \omega_4 = \text{上升,上升}\}$$

通过取至少两个时间段，我们可以对随机利率 r_t 的影响进行研究．图 17-1 和图 17-2 描述了这一过程．若投资者希望在 t_1 到 t_2 期间借款，他可以在 t_1 开始执行无风险利率合约．那么无论将来发生什么情况，他的回报都是无风险的，因为回报是已经确定的．[⊖]

不管 t_2 发生的状态是上升还是下降，投资者获得的收入都是$\{1 + r_{t_1}\Delta\}$. 由于无论发生什么状态，无风险借贷款在 t_2 时刻的收益相同，因此在模型针对 t_1 到 t_2 的部分中，可以看作即期利率无波动，所以无须分析 r_t 随机性的影响．

但当我们增加时间区间的数量，情况就会发生变化．$t_1 = 1$ 时刻提供给投资者的即期利率到了 $t_2 = 1 + \Delta$ 会面临一些风险．在这种情况下，投资者可能不再借出自己的资金，无论是以

⊖ 假设没有违约风险．

较高的利率 $r_{t_2}^u$ 还是以较低的利率 $r_{t_2}^d$. t_2 时刻的即期利率在 t_1 时刻是未知的．所以，在有 3 个时间点 t_1，t_2，t_3 的情况下，[一]“无风险”利率的随机性将会成为一个重要因素，尽管在第 2 章单阶段的框架中这并不重要．在时间点 t_2 有两个可能的即期利率，$r_{t_2}^u$ 和 $r_{t_2}^d$，故 r_{t_2} 是随机的．[二]

280

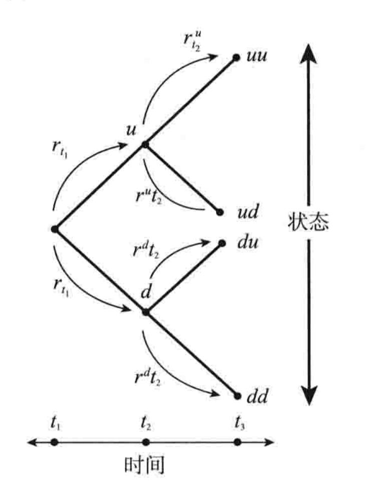


图 17-1 无风险即期利率的随机图示

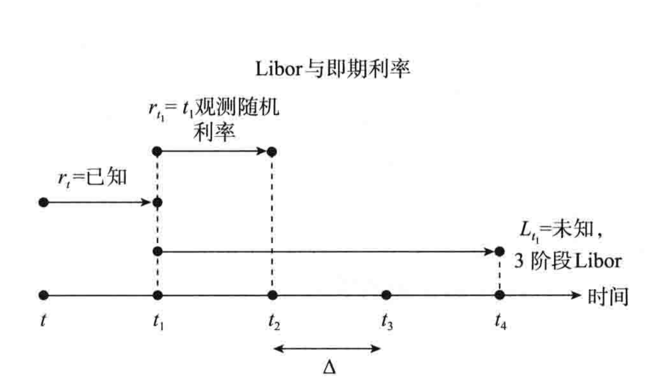


图 17-2 随机即期利率 r_t 的影响

和第 2 章相比，另一个不同之处是我们对于所研究的金融工具的选择．之前我们讨论股票和建立在股票上的期权，现在我们研究利率敏感型证券和远期合约．这样我们可以将无套利价格应用到别的有趣的资产上，并且直观地展示出改变标准化方法对于金融市场参与者而言是一种有效的工具．特别地，在这种框架下我们能够引入所谓的远期测度，并将它的性质与风险中性测度比较．

所以我们假设，现有以下金融工具的流动市场：[三]

- 无违约风险的储蓄账户．在时间 t 投资者可以以利率 r_t 为合约，经过时间 Δ 后，该投资者将收到$(1+r\Delta)$. 如果投资者想要继续短期投资，那么它将在 $t+\Delta$ 时刻以新的利率$r_t+\Delta$ 投资．
- 利率为 L_t 的远期合约．该利率是无违约风险的．可以关于此利率建立超过一个时间段的合约．例如，它可以作为为期 6 个月的商业借款的简单利率，比如 6 个月 Libor. 所以我们选择符号 L_t. [四]

[一] 这意味着有两个时间段．

[二] 从这里可以看出，对储蓄(或货币市场)账户而言，用“无风险”来形容并不是完全恰当的．投资是无风险的，意味着在合约到期时我们知道能获得多少收益．这种合约没有“市场”风险．其价格在合约期间不会波动，因为 r_t 是常数．此外，我们假设违约风险也不存在．所以合约结束时的收益是常数．但是，如果这项投资是滚动的，投资者储蓄账户内资金的收益会发生随机波动．

[三] 在这里，市场的流动性表示资产可以随时按估价买卖．

[四] 和我们上一章一样，Libor 是伦敦同业拆借利率，伦敦的银行可以以此利率借款．Libor 利率常被用作基准利率，建立在 Libor 基础上的金融工具是银行资产负债表上资产的重要组成部分．

- t 时刻价格为 $B(t,t_3)(t<t_3)$ 的无违约风险短期折现债券. 这种债券在到期日支付一美元，其他时期不支付.
- t 时刻价格为 $B(t, T)$ 的无违的风险长期折现债券. 这种债券在到期日 T 支付票面金额，设为 1. 由于我们想让债券的到期日在所研究的时间段之后，令 $t<t_3<T$. 此外我们没有必要为 T 取具体数值.
- 建立在 L_{t_2} 上的远期利率协议(FRA). 其中 F_{t_2} 是在 t_1 时刻约定的远期利率. 如果 $F_{t_2}>L_{t_2}$，FRA 的购买者将支付差额. 要注意，在 t_2 时刻，支付额所依赖的利率即为已知，但要到 t_3 时刻才实施相应支付. 这造成了 FRA 在支付上的延迟. Δ 是根据天数调整的时间.

281

- 最后，我们考虑利率衍生品，比如关于 $B(t_1,t_3)$ 的看涨期权，或者关于 L_{t_2} 的上限协议. 该衍生品在 $t=t_3$ 到期，且现价为 C_{t_1}.

现在我们将这些资产及对应的回报放在矩阵方程中，与第 2 章类似. 但是，首先我们需要进行符号简化.

从 t_1 到 t_2 的总收入为:

$$R_{t_1} = (1+r_{t_1}\Delta)$$
$$R_{t_2} = (1+r_{t_2}\Delta)$$

在之后的几章，我们将使用原始标记. 不过这里为了简化矩阵方程，我们还是会对债券使用简化的标记. 令

$$B^s_{t_1} = B(t_1,t_3)$$

它表示短期债券在 t_1 时刻的价格. 而

$$B_{t_1} = B(t_1,T)$$
$$B_{t_3} = B(t_3,T)$$

分别表示长期债券在 t_1 和 t_3 的价格. 我们不需要考虑长期债券在 t_2 的价格.

我们令名义金额为 N 的 FRA 记作金额为 1，因为该参数对于我们模型的结果没有影响.

最后，我们约定所有利率都是以 Δ 为时间单位的利率，这样我们就不需要为了计算收益而用 Δ 乘以年利率，而且还能起到简化记号的效果. 或者，我们也可以认为 $\Delta=1$ 等于一年.

现在我们可以写出金融基本定理所对应的矩阵方程. 将上面罗列的 5 种工具的价格以 (5×1) 向量的形式放在左边，可以得到以下关系:

$$\begin{bmatrix} 1 \\ 0 \\ B^s_{t_1} \\ B_{t_1} \\ C_{t_1} \end{bmatrix} = \begin{bmatrix} R_{t_1}R^u_{t_2} & R_{t_1}R^u_{t_2} & R_{t_1}R^d_{t_2} & R_{t_1}R^d_{t_2} \\ (F_{t_1}-L^u_{t_2}) & (F_{t_1}-L^u_{t_2}) & (F_{t_1}-L^d_{t_2}) & (F_{t_1}-L^d_{t_2}) \\ 1 & 1 & 1 & 1 \\ B^{uu}_{t_3} & B^{ud}_{t_3} & B^{du}_{t_3} & B^{dd}_{t_3} \\ C^{uu}_{t_3} & C^{ud}_{t_3} & C^{du}_{t_3} & C^{dd}_{t_3} \end{bmatrix} \times \begin{bmatrix} \psi^{uu} \\ \psi^{ud} \\ \psi^{du} \\ \psi^{dd} \end{bmatrix} \tag{17.6}$$

其中右端第一部分为各产品在 t_3 可能的收益，第二部分为状态价格 ψ_i. 该矩阵方程与第 2 章中的类似，下面说一些需要注意的地方．

该系统的第一行描述了无风险储蓄投资的可能结果．若投资 1 美元，t_2 的回报是已知的 $R_{t_1}=(1+r_{t_1})$，t_3 的回报是未确定的 $R_{t_2}=(1+r_{t_2})$. t_3 的收益具有随机性，因为和 R_{t_1} 不同，在 t_1 时刻还不能确定 R_{t_2} 的值．在 t_2 有两种可能性，由上标 u、d 记作 $R_{t_2}^u$、$R_{t_2}^d$. 该行与(17.1)的第一行相似，但是这里的元素不是常数．

其次，考虑矩阵方程的第二行．F_{t_1} 是关于随机的 Libor 利率 L_{t_2} 在 t_1 时刻约定的远期利率，而 L_{t_2} 在 t_2 时刻才能观察到．所以这是一个远期利率合约．FRA 在签订时的无套利价值为 0，签订这类合约不需要预付款．这也解释了左边向量第二个元素为什么是 0. 根据 FRA 合约，已知的 F_{t_1} 和未知的 L_{t_3} 之差将在 t_3 时刻支付，这解释了矩阵的第二行．有两种不同的结果． 282

矩阵方程的第三、四行代表两种债券．$B_{t_1}^s$ 和 B_{t_1} 都表示零息债券在 t_1 时刻的无套利价格，前者的到期日为 t_3，后者为某个未来的时刻 T. 注意在右边矩阵中，期限较短的债券在 t_3 价值为常数 1，因为此时该债券到期，回报为 1. 而长期债券没有这种性质，B_{t_3} 有 4 种可能的取值．

最后，$\{\psi^{ij}, i, j=u, d\}$ 为 t_3 时刻的 4 种状态价格．当且仅当不存在套利机会时，它们取正值．与第 2 章相同，对于所有 i，j，必须满足：

$$\psi^{ud}>0$$

考虑现在的情况与第 2 章有哪些区别．首先，在该矩阵方程中，我们可以清楚地观察到“无风险”储蓄投资的风险．投资仅在第一个时间段是无风险的．投资者确定自己在下一个时间段即 $t_2=t+\Delta$ 时刻的收益．但是，再往后一个时间段，即期利率可能会更高或更低，这取决于状态 $\omega_{i,j}$. 所以通常来说，即期利率 r_t 是已知的，而 $r_{t+\Delta}$ 是随机的，可能升高也可能降低．所以它带有上标，注明价格对应的状态．

其次，该模型中的证券与第 2 章有很大不同．远期合约和债券都是利率敏感型工具，这些工具的定价更为复杂，而第 2 章所选资产的定价模型简单得多．对于期权 C_t 而言亦是如此，此时该期权建立在利率敏感型证券上．

最后，注意该模型还有直观的一面．由于其中一支债券在 t_3 到期，它在 t_3 的收益是已知的且为常数．这一点对于综合概率(鞅测度)的选择非常重要，而利用综合概率能够使这些新型资产的定价更加方便．

现在我们可以开始考虑标准化的问题，进而选择测度．但首先我们需要作出以下说明．

要点

本章中，R_t 表示短期利率，L_t 表示 Libor 过程．从本质来看，它们是两个不同的过程. L_t 的到期时间与即期利率不同．根据定义，即期利率的到期时间为距当前最近的下一时刻．但是为了让模型及所研究的金融工具尽量简化，我们假设 L_t 是一个单位时间上的 Libor. 这样 R_t 和 L_t 将会相等．不过即使这样设定，本章的结论仍然是成立的．本章从名

义上区分了这两种利率，因为如果 Libor 利率的期限延长，这种等价关系将不存在．对 L_t 和 R_t 用不同的记号，能够帮助我们理解 Libor 工具以及在更一般情况下它们之间的关系．

另一种方法是考虑两个阶段的 Libor. 但这需要三阶段的模型，将会使矩阵方程比现在的更为复杂．因此双阶段、四种状态的模型是最小的系统，我们将讨论这种情况下的标准化.

17.2.2 标准化

283

我们还是从回顾第 2 章的框架开始．回想当时我们是如何通过金融基本定理得到风险中性概率的．为了更具体地说明，我们写出(17.1)以及之前第 2 章中关于 S_t 的方程：

$$S_t = S^u_{t+\Delta}\psi^u + S^d_{t+\Delta}\psi^d \tag{17.7}$$

为了在该方程中引入风险中性概率，我们对每个 ψ^i 乘以$(1+r\Delta)/(1+r\Delta)$：

$$S_t = S^u_{t+\Delta}\psi^u \frac{(1+r\Delta)}{(1+r\Delta)} + S^d_{t+\Delta}\psi^d \frac{(1+r\Delta)}{(1+r\Delta)} \tag{17.8}$$

然后可以看到，$\psi^i(1+r\Delta)$实际上就是 Q_i. 那么可得

$$S_t = \frac{1}{(1+r\Delta)}S^u_{t+\Delta}Q_u + \frac{1}{(1+r\Delta)}S^d_{t+\Delta}Q_d \tag{17.9}$$

当我们用 Q_i 代替 $\psi^i(1+r\Delta)$之后，每一项增加了一个分子为 1，分母为$(1+r\Delta)$的因子. $(1+r\Delta)$表示该情况下无风险投资的回报．第一阶段的无风险回报是常数，因此可以将该因子提出来，得到：

$$S_t = \frac{1}{(1+r\Delta)}[S^u_{t+\Delta}Q_u + S^d_{t+\Delta}Q_d] \tag{17.10}$$

但是，在新模型中有两个阶段，该回报是随机变量，相同的推导在这里无法进行．

而且还要记住，在定价方程中引入风险中性测度，可以将各状态的收益标准化为对应的无风险贷款的收益．用 Q_i 替代 ψ^i，等价于将矩阵(17.1)中每行的元素都除以第一行对应的元素，而第一行的元素代表了该状态下无风险投资的收益．

正如之前所说，在风险中性测度 Q 下，资产价格具有趋势．实际上，在 Q 下，所有期望收益率都等于 r，这意味着价格将上移．所以这些价格在 Q 下不是鞅．但是，用无风险利率标准化之后，期望收益率(代表了趋势)将变为 0. 也就是说，分子和分母都将以 r 的速率增加，标准化后的变量将成为鞅，价格不再有明显的趋势．

注意这种标准化的另一个性质．除以$(1+r\Delta)$相当于折现一笔未来的现金流．但是，在用无风险贷款进行标准化时，我们会先折现，然后用概率求平均，得到：

$$S_t = E^Q\left[\frac{S_{t+\Delta}}{(1+r\Delta)}\right] \tag{17.11}$$

当 r 为常数时，该式可化简为

$$S_t = \frac{1}{(1+r\Delta)}E^Q[S_{t+\Delta}] \tag{17.12}$$

这就是我们在本书第一部分使用过多次的定价方程．与本章所有期望算子一样，除非特别

说明，E^Q 表示以 t 时刻信息为条件的条件期望．

以上步骤在利率衍生品的双阶段模型中是否还能成立呢？答案是否定的．

尝试在(17.6)的新模型中采用同样策略．假定我们还是通过系统中对应储蓄账户的第一行来得到风险中性概率 Q^{ij}：

$$1 = R_{t_1} R_{t_2}^u \psi^{uu} + R_{t_1} R_{t_2}^u \psi^{ud} + R_{t_1} R_{t_2}^d \psi^{du} + R_{t_1} R_{t_2}^d \psi^{dd} \tag{17.13}$$

其中只有 R_{t_2} 有状态上标，因为在 t_1 时刻，R_{t_1} 是已经确定的．

我们可以用与之前类似的方法定义 4 个风险中性概率：

$$Q_{ij} = (1 + r_{t_1})(1 + r_{t_2}^i)\psi^{ij} \tag{17.14}$$

284

其中 i，$j=u$，d. 方程(17.13)即成为：

$$1 = Q_{uu} + Q_{ud} + Q_{du} + Q_{dd} \tag{17.15}$$

如果不存在套利机会，状态价格$\{\psi^{ij}\}$为正，且我们有

$$Q_{ij} > 0 \tag{17.16}$$

那么，像第 2 章一样，我们可以把 Q^{ij} 当作各状态对应的概率，即便它们不表示这 4 种状态在现实世界中的概率．按第 2 章的方法继续这一推导，我们可以利用系统(17.6)中剩下的方程得到对应的鞅等式．

例如，系统的第三行给出了测度 Q 的下短期债券无套利价格：

$$B_{t_1}^s = \frac{1}{(1 + r_{t_1})(1 + r_{t_2}^u)} Q_{uu} + \frac{1}{(1 + r_{t_1})(1 + r_{t_2}^u)} Q_{ud} + \frac{1}{(1 + r_{t_1})(1 + r_{t_2}^d)} Q_{du} + \frac{1}{(1 + r_{t_1})(1 + r_{t_2}^d)} Q_{dd} \tag{17.17}$$

或记为

$$B_{t_1}^s = E^Q\left[\frac{1}{(1 + r_{t_1})(1 + r_{t_2})}\right] \tag{17.18}$$

其中 r_{t_2} 是随机的，不能提到(条件)期望算子外面．

推广到连续时间并且假定连续状态，设到期时间为 $T(t<T)$，我们可以将该公式一般化．无违约风险的零息债券的无套利价格可以写为

$$B(t,T) = E^Q[\mathrm{e}^{-\int_t^T r_u \mathrm{d}u}] \tag{17.19}$$

该方程我们会在之后的章节进一步使用．

现在我们可以得到系统(17.6)中的长期债券类似的定价公式，步骤与之前相同．(17.6)的第四行可以得到：

$$B_{t_1} = \frac{B_{t_3}^{uu}}{(1 + r_{t_1})(1 + r_{t_2}^u)} Q_{uu} + \frac{B_{t_3}^{ud}}{(1 + r_{t_1})(1 + r_{t_2}^u)} Q_{ud} + \frac{B_{t_3}^{du}}{(1 + r_{t_1})(1 + r_{t_2}^d)} Q_{du} + \frac{B_{t_3}^{dd}}{(1 + r_{t_1})(1 + r_{t_2}^d)} Q_{dd} \tag{17.20}$$

或记为

$$B^s_{t_1} = E^Q\left[\frac{B_{t_3}}{(1+r_{t_1})(1+r_{t_2})}\right] \tag{17.21}$$

这里 r_{t_2} 也是一个随机变量．而 r_{t_3} 也是随机的，因为 t_3 并不是该债券的到期日．由于这个原因，这种方程在实际定价中不是非常有用．

最后，利用(17.6)的第二行和第五行，可以分别得到两种 Libor 工具的定价方程，即 FRA 和上限协议衍生品 C_t：

$$0 = E^Q\left[\frac{1}{(1+r_{t_1})(1+r_{t_2})}[F_{t_1} - L_{t_2}]\right] \tag{17.22}$$

$$C_{t_1} = E^Q\left[\frac{1}{(1+r_{t_1})(1+r_{t_2})}C_{t_3}\right] \tag{17.23}$$

因此，当我们采用第 2 章中的方法，用储蓄账户来定义 Q，确实可以得到类似于(17.18)或 285 (17.19)的定价公式．但是，对于我们现在研究的这些具有随机利率的金融工具而言，使用这种风险中性概率 Q 并不能带来方便，有时甚至是不合适的．它会给市场参与者带来不必要的困难．下一节我们将具体介绍．

17.2.3 一些不良性质

概率 Q^{ij} 是通过储蓄账户方程得到的：

$$1 = R_{t_1}R^u_{t_2}\psi^{uu} + R_{t_1}R^u_{t_2}\psi^{ud} + R_{t_1}R^d_{t_2}\psi^{du} + R_{t_1}R^d_{t_2}\psi^{dd} \tag{17.24}$$

将 Q^{ij} 带入后即为

$$1 = Q_{uu} + Q_{ud} + Q_{du} + Q_{dd} \tag{17.25}$$

现在考虑这些概率在 FRA 合约定价中具体如何运用．首先意识到，对 FRA 合约定价意味着找到合适的 F_{t_i}，使得该合约在 t_1 时刻价值为 0. 因为所有 FRA 在开始刚交易的时候价格为 0，且该价格为无套利价格．我们的任务就是找到对应的无套利的 F_{t_i}．从系统(17.6)的第二行可以得到

$$0 = (F_{t_1} - L^u_{t_2})\psi^{uu} + (F_{t_1} - L^u_{t_2})\psi^{ud} + (F_{t_1} - L^d_{t_2})\psi^{du} + (F_{t_1} - L^d_{t_2})\psi^{dd} \tag{17.26}$$

对右边各项乘并除以$(1+r_{t_1})(1+r^i_{t_2})$，并利用

$$(1+r_{t_1})(1+r^i_{t_2})\psi^{ij} = Q^{ij} \tag{17.27}$$

重新组合，可得

$$\begin{aligned}0 = &\frac{(F_{t_1} - L^u_{t_2})}{(1+r_{t_1})(1+r^u_{t_2})}Q^{uu} + \frac{(F_{t_1} - L^u_{t_2})}{(1+r_{t_1})(1+r^u_{t_2})}Q^{ud} \\ &+ \frac{(F_{t_1} - L^d_{t_2})}{(1+r_{t_1})(1+r^d_{t_2})}Q^{du} + \frac{(F_{t_1} - L^d_{t_2})}{(1+r_{t_1})(1+r^d_{t_2})}Q^{dd}\end{aligned} \tag{17.28}$$

F_{t_1} 的值与未来发生哪种状态无关，将它提到括号外面：

$$F_{t_1}\left[\frac{Q^{uu}}{(1+r_{t_1})(1+r^u_{t_2})} + \frac{Q^{ud}}{(1+r_{t_1})(1+r^u_{t_2})} + \frac{Q^{du}}{(1+r_{t_1})(1+r^d_{t_2})} + \frac{Q^{dd}}{(1+r_{t_1})(1+r^d_{t_2})}\right]$$

$$= \frac{L_{t_2}^u Q^{uu}}{(1+r_{t_1})(1+r_{t_2}^u)} + \frac{L_{t_2}^u Q^{ud}}{(1+r_{t_1})(1+r_{t_2}^u)} + \frac{L_{t_2}^d Q^{du}}{(1+r_{t_1})(1+r_{t_2}^d)} + \frac{L_{t_2}^d Q^{dd}}{(1+r_{t_1})(1+r_{t_2}^d)} \tag{17.29}$$

这样我们可以写出：

$$F_{t_1} E^Q\left[\frac{1}{(1+r_{t_1})(1+r_{t_2})}\right] = E^Q\left[\frac{L_{t_2}}{(1+r_{t_1})(1+r_{t_2})}\right] \tag{17.30}$$

重新组合，可以得到无套利 FRA 利率的定价公式：

$$F_{t_1} = \frac{1}{E^Q\left[\frac{1}{(1+r_{t_1})(1+r_{t_2})}\right]} E^Q\left[\frac{L_{t_2}}{(1+r_{t_1})(1+r_{t_2})}\right] \tag{17.31}$$

该公式中，我们利用无风险概率 Q 来确定合约利率 F_t. 但是，与常利率的期权定价不同，在该方法中我们很快就会发现一些影响我们推导的性质．[286]

首先，一般来说，F_t 不是 L_{t_2} 的无偏估计：

$$F_{t_1} \neq E^Q[L_{t_2}] \tag{17.32}$$

除非 r_t 和 L_t 是统计独立的．如果独立，那么就可以根据(17.31)分别求期望：

$$F_{t_1} = \frac{1}{E^Q\left[\frac{1}{(1+r_{t_1})(1+r_{t_2})}\right]} \times E^Q\left[\frac{1}{(1+r_{t_1})(1+r_{t_2})}\right] E^Q[L_{t_2}] \tag{17.33}$$

化简，即得

$$F_{t_1} = E^Q[L_{t_2}] \tag{17.34}$$

在这种苛刻假设下，远期利率才能作为其对应 Libor 过程的无偏估计．但是，在现实中，我们能否说短期利率和期限更长的 Libor 利率是统计独立的呢？这是一个很难成立的假设．

第二，考虑使用风险中性测度 Q 的另一个弊端．之前我们已经注意到，在 Q 测度下求期望时，含有即期利率的项不能够提到期望算子的外面．而在第 2 章的简单模型中，r 是与状态无关的常数．而现在 r_{t_2} 的值与状态 u，d 有关．所以，在方程(17.31)分母中，期望算子内的项是随机的，且只能放在求期望的符号内．

第三，(17.31)中 FRA 的定价公式是非线性的．这个性质初看起来没什么不利的影响，但在使用风险中性测度的时候会成为很大的缺陷．它将为市场参与者带来非常大的不便．事实上，当我们试图确定 FRA 利率 F_{t_1} 或者衍生品价格 C_t 时，我们现在需要对两个过程建模，即 r_t 和 L_t，而非 L_t 一个过程．更糟糕的是，这两个过程有着复杂的相关性．对这样的非线性表达式求期望，是一项艰巨的任务．

最后，根据定义，与 F_{t_1} 相关的现金到 t_3 时刻才到期．现考虑风险中性测度在定价方程(17.31)中如何运用．在此定价公式中，首先对 t_3 时刻的现值进行折现，然后通过期望运算求平均，再乘以折现项期望的倒数．这样才符合合约的意义．

显然，这并不是计算无套利远期利率的有效方式．实际上，我们可以将这种折现去掉，因为我们都是在 t_3 时刻比较 F_{t_1} 和 L_{t_2} 的值并结算．

以同样的方法推导 Libor 衍生品 C_t，可以得到同样的结论．定价方程为：

$$C_{t_1}=E^Q\left[\frac{1}{(1+r_{t_1})(1+r_{t_2})}C_{t_3}\right] \tag{17.35}$$

同样，如果 C_t 是利率敏感型衍生品，那么存在同样的关于 Q 的问题．随机的折现因子不能提到期望算子的外面，而且如果 C_{t_3} 是利率敏感型证券，即期利率很可能与 C_{t_3} 是统计相关的.

所以，如果按照方程(17.24)定义概率测度 Q，就会产生第 2 章中没有的问题．下面我们将介绍一种更好的方法来选取综合概率，从而方便有效地避免上述问题．

17.2.4 新的标准化方法

287

现在我们考虑另一种得到鞅概率的方法．同样以(17.6)作为开端，根据第三行我们可以写出：

$$B^s_{t_1}=\psi^{uu}+\psi^{ud}+\psi^{du}+\psi^{dd} \tag{17.36}$$

除以 $B^s_{t_1}$，得

$$1=\frac{1}{B^s_{t_1}}\psi^{uu}+\frac{1}{B^s_{t_1}}\psi^{ud}+\frac{1}{B^s_{t_1}}\psi^{du}+\frac{1}{B^s_{t_1}}\psi^{dd} \tag{17.37}$$

取

$$\pi^{ij}=\frac{1}{B^s_{t_1}}\psi^{ij} \tag{17.38}$$

方程可改写为

$$1=\pi_{uu}+\pi_{ud}+\pi_{du}+\pi_{dd} \tag{17.39}$$

由于在无套利条件下，ψ^{ij} 大于 0，故

$$\pi_{ij}>0,\quad i,j=u,d \tag{17.40}$$

这意味着 π_{ij} 可以成为一种新的综合鞅概率．它们可以得出一组新的鞅关系．我们将 π_{ij} 称为远期测度．在考虑远期测度 π_{ij} 相比于风险中性测度 Q_{ij} 的优势之前，我们先对这一新的标准化方法作几点说明．

首先，要从状态价格 ψ^{ij} 得到 π，我们需要对每一个状态价格乘以 $B^S_{t_1}$，后者是在 t_1 时刻确定的值，与未来状态无关，不带有状态上标．这意味着在用 π 求期望时，$B^S_{t_1}$ 是可以移到期望符号外面的．

第二，对于不同期限的无违约风险零息债券，我们都可以定义出新的远期测度．因此，可能需要在测度上加一个时间下标，如 π_T，表示对应债券的到期时间为 T. 由于衍生品建立在利率敏感型证券上，使用与衍生品同时到期的债券来建立远期测度，显然更为合适．

最后，注意观察标准化是如何完成的．为了在定价方程中引入这种概率，我们先对每个 ψ^{ij} 乘同时除以 $B^S_{t_1}$. 用 π_{ij} 替代$\frac{\psi^{ij}}{B^S_{t_1}}$之后，相当于对方程(17.6)的矩阵中每一个资产价格都乘以短期债券价格 $B^S_{t_1}$. 所以，可以说我们是用 $B^S_{t_1}$ 进行标准化．

相关内容将在下面具体介绍．

标准化的性质

现在我们讨论一些使用 π 代替 Q 作为概率测度的重要结果．

下面我们一步一步介绍．首先，当使用风险中性测度时，我们会发现 F_{t_1} 是 Libor 过程 L_t 的有偏估计．也就是说

$$F_{t_1} \neq E^Q[L_{t_2}] \tag{17.41}$$

现在考虑在测度 π 下求类似的期望．我们先取系统(17.6)中的第二行，对每个元素乘以 $B_{t_1}^S / B_{t_1}^S$：

$$\begin{aligned} 0 = & (F_{t_1} - L_{t_2}^u) \frac{B_{t_1}^s}{B_{t_1}^s} \psi^{uu} + (F_{t_1} - L_{t_2}^u) \frac{B_{t_1}^s}{B_{t_1}^s} \psi^{ud} \\ & + (F_{t_1} - L_{t_2}^d) \frac{B_{t_1}^s}{B_{t_1}^s} \psi^{du} + (F_{t_1} - L_{t_2}^d) \frac{B_{t_1}^s}{B_{t_1}^s} \psi^{dd} \end{aligned} \tag{17.42}$$

由于比率

$$\frac{\psi^{ij}}{B_{t_1}^s} \tag{17.43}$$

就是对应的 π_{ij}，且它们的和为 1，我们可以化简并得到：

288

$$0 = B_{t_1}^s [F_{t_1} - [L_{t_2}^u \pi_{uu} + L_{t_2}^u \pi_{ud} + L_{t_2}^d \pi_{du} + L_{t_2}^d \pi_{dd}]] \tag{17.44}$$

注意到 $B_{t_1}^S$ 可以被提取到了括号外面，因为它是常数，根据观察可以得到无套利价格 $B_{t_1}^S$．因此，

$$F_{t_1} = [L_{t_2}^u \pi_{uu} + L_{t_2}^u \pi_{ud} + L_{t_2}^d \pi_{du} + L_{t_2}^d \pi_{dd}] \tag{17.45}$$

很容易发现，右端是 Libor 过程在新的鞅概率 π_{ij} 下的期望值．这意味着现在我们有：

$$F_{t_1} = E^\pi[L_{t_2}] \tag{17.46}$$

于是我们得到了一个重要结论．虽然在经典的风险中性测度下，F_{t_1} 是 L_{t_2} 的有偏估计量，但是在远期测度 π 下，它是 L_{t_2} 的无偏估计量．

为什么这一点很重要？它有着怎样的实际应用？

考虑以下一般情况，我们用 Δ 来代替 t_i 作为时间记号．设 $t+2\Delta$ 时刻的 Libor 利率为 $L_{t+2\Delta}$. 当前的远期利率为 F_t，考虑它的未来值 $F_{t+\Delta}$，$\Delta > 0$. ⊖ 我们可以利用测度 π 并写出：

$$F_{t+\Delta} = E_{t+\Delta}^\pi [L_{t+2\Delta}] \tag{17.47}$$

其中 $E_t^\pi[\cdot]$的下标表示该期望是以 t 时刻的已知信息为条件的条件期望．也就是说，

$$E_t^\pi[\cdot] = E^\pi[\cdot \mid I_t]$$

其中 I_t 是 t 时刻可获得的信息集．在本例中，它包括系统中所有资产当前和过去的价格．

之前我们介绍过条件期望运算的递归性：

⊖ 所以 F_t 是“当前”观察到的 FRA 利率，而 $F_{t+\Delta}$是经过一小段时间 Δ 后观察到的利率．

$$E_t^\pi[E_{t+\Delta}^\pi[\bullet]] = E_t^\pi[\bullet] \tag{17.48}$$

该性质表示，要得到当前对于未来某时刻预测结果的预测，“最好”的方法就是依据当前信息预测.[⊖]

由于在测度 π 下，F_t 是 $L_{t+2\Delta}$ 的无偏估计量，我们可以写出：

$$F_t = E_t^\pi[L_{t+2\Delta}] \tag{17.49}$$

利用条件期望递归性，我们可以在适当的位置引入 $E_{t+\Delta}^\pi$ 算子：

$$F_{t+\Delta} = E_t^\pi[E_{t+\Delta}^\pi[L_{t+2\Delta}]] \tag{17.50}$$

将(17.47)代进来，可得：

$$F_t = E_t^\pi[F_{t+\Delta}] \tag{17.51}$$

这说明过程$\{F_t\}$在远期测度 π 下是鞅．后面我们会看到，远期价格的这种性质在对某些利率敏感型金融工具的定价中非常方便．从衍生品 C_t 的条件期望中就可以看出这一点．

假定 C_t 是一个上限协议的价格．在到期日，该上限协议将支付：

$$C_{t_3} = N \max[L_{t_2} - K, 0] \tag{17.52}$$

其中 N 是名义金额，我们令它为 1. K 是 t_1 时刻选取的上限利率，L_{t_2} 是 t_2 时刻的 Libor 利率. 支付将拖延到 t_3 时刻完成．因此该合约能够给购买者一种保险，确保他们以不超过 K 的利率借款．

[289] 我们该怎样对这种金融工具定价呢？考虑使用经典风险中性测度 Q. 根据风险中性测度 Q 我们有：

$$C_{t_1} = E^Q\left[\frac{1}{(1+r_{t_1})(1+r_{t_2})}\max[L_{t_2} - K, 0]\right] \tag{17.53}$$

根据之前的讨论，在该定价方程中，(随机)即期利率 r_{t_2} 一般与(随机)Libor 利率 L_{t_2} 相关．所以如果市场参与者要用这种方法对上限协议定价，必须对双变量过程 r_t，L_t 建模．

但是，如果对系统(17.6)的最后一个方程使用远期测度，将得到

$$C_{t_1} = B_{t_1}^s E^\pi[C_{t_3}] \tag{17.54}$$

这意味着

$$C_{t_1} = B_{t_1}^s E^\pi[\max[L_{t_2} - K, 0]] \tag{17.55}$$

在该方程中，我们用短期债券的无套利价格乘以 L_{t_2} 的函数的条件期望．也就是说，我们完全省去了对双变量过程建模的问题．期望算子内只有一个随机变量 L_{t_2}.

我们可以看到新的测度能够带来以下便利．在远期测度 π 下，我们首先计算 t_2 时刻(远期)的期望，然后用观察到的无套利价格 $B_{t_1}^S$ 进行折现．而在风险中性测度中，我们先对随机的收益使用一个随机的折现因子，然后再用概率求平均．我们使用风险中性测度这样做的时候，错过了使用市场提供的折现因子的机会，即 $B_{t_1}^S$. 相反地，我们需要计算折现因子，需要涉及复杂的二元过程．下一节我们还将介绍一个例子．

⊖ 这之前一样，“最好”是针对均方误差而言的．

17.3 其他等价鞅测度

我们已经介绍了远期测度．本节我们还将涉及以下测度：股份测度，互换测度，即期测度．

17.3.1 股份测度

计价单位是 S_t 的测度叫做股份测度，S．任何可交易证券除以 S_t 后就会成为股份测度下的鞅．这意味着

$$\frac{C_t}{S_t}=E_t^S\left(\frac{C_T}{S_T}\right) \tag{17.56}$$

对于执行价格为 K 的看涨期权，我们可以写出

$$\frac{C(K)}{S_0}=E^S\left(\frac{(S_T-K)^+}{S_T}\right) \tag{17.57}$$

如果我们假设证券价格 S_T 恒为正，我们可以得到

$$\frac{C(K)}{S_0}=E^S\left(\left(1-\frac{K}{S_T}\right)^+\right) \tag{17.58}$$

我们定义 $y=\log\left(\frac{S_T}{K}\right)$，即意味着$\frac{K}{S_T}=\mathrm{e}^{-y}$．利用这种定义，我们可以将标准化的看涨期权价格改写为：

$$\frac{C(K)}{S_0}=\int_0^{\infty}(1-\mathrm{e}^{-y})f(y)\mathrm{d}y \tag{17.59}$$

其中 $f(y)$是在股权测度下的概率密度函数，$y=\ln(S/K)$．通过分部积分可以得到 290

$$\begin{aligned}\frac{C(K)}{S_0}&=\int_0^{\infty}(1-\mathrm{e}^{-y})f(y)\mathrm{d}y=(1-\mathrm{e}^{-y})F(y)\bigg|_0^{\infty}-\int_0^{\infty}\mathrm{e}^{-y}F(y)\mathrm{d}y\\&=(F(y)-F(y)\mathrm{e}^{-y})\bigg|_0^{\infty}-\int_0^{\infty}\mathrm{e}^{-y}F(y)\mathrm{d}y=1-\int_0^{\infty}\mathrm{e}^{-y}F(y)\mathrm{d}y\\&=\int_0^{\infty}\mathrm{e}^{-y}\mathrm{d}y-\int_0^{\infty}\mathrm{e}^{-y}F(y)\mathrm{d}y=\int_0^{\infty}(1-F(y))\mathrm{e}^{-y}\mathrm{d}y\end{aligned}$$

因此我们有

$$\frac{C(K)}{S_0}=\int_0^{\infty}(1-F(y))\mathrm{e}^{-y}\mathrm{d}y$$

对于给定的 y，表达式 $1-F(y)$对应 $\ln(S/K)$大于 y 的概率．将 e^{-y}作为一个 $\lambda=1$ 的指数型随机变量的概率密度函数，标准化后的看涨期权价格是在股份测度下，股票价格对数大于执行价格对数且差值为一个独立的指数型变量的概率(Carr 和 Madan(2009))．或者等价地，

$$\frac{C(K)}{S_0}=P(\ln(S/K)>Y) \tag{17.60}$$

$$= P(\ln S - \ln K > Y) \tag{17.61}$$
$$= P(X - Y > \ln K) \tag{17.62}$$

其中 X 是在股份测度下股票价格的对数，Y 是一个独立的指数型变量，K 是执行价格．

17.3.2 即期测度和市场模型

远期利率合约(FRA)是一种可以用于直接交易简单远期利率的可交易合同．该合同包含 3 个时间点：(a)当前时刻 t，(b)失效时刻 T，(c)到期时刻 S，$S>T$. 在 S 时刻合约的支付额是 $1+(S-T)F(t;T,S)$，它将造成 T 时刻一美元的远期投资．但是这种投资可以用以下策略复制：

在 t 时刻卖出面值为 1 的 T 时刻到期的债券 T-bond，买入$\frac{P(t,T)}{P(t,S)}$的到期日为 S 的债券S-bond. 该投资花费为 0.

在 T 时刻，T-bond 到期，我们要向债券的购买者支付 1 美元．

在 S 时刻，S-bond 到期，我们将收到$\frac{P(t,T)}{P(t,S)}$美元．

以上交易的实际结果，是造成 T 时刻 1 美元的远期投资和 S 时刻$\frac{P(t,T)}{P(t,S)}$美元的收入．根据无套利条件，有以下方程，$1+(S-T)F(t;T,S)=\frac{P(t,T)}{P(t,S)}$，它意味着

$$F(t;T,S) = \frac{1}{S-T}\left(\frac{p(t,T)}{P(t,S)} - 1\right) \tag{17.63}$$

这叫做$[T,S]$在 t 的简单(简单复合)远期利率．有固定间隔 Δ 的简单远期利率叫做 Libor 利率，$L(t,T)=F(t,T,T+\Delta)$，其中 Δ 通常取特定的值，如$\frac{1}{4}$或$\frac{1}{2}$.

在市场模型中，Libor 利率或互换利率往往被作为基础利率．但是如果选择这些利率的其中一种，并且使用风险中性测度来进行定价，我们会发现这其实非常不便利．因为定价需要确定各时间点瞬时的短期利率．因此我们通常需要用到其他测度来进行定价．其中一种就是即期测度．我们首先选取固定的适用于市场模型的到期时间和间隔天数．考虑流动性证券的价格，如上限/下限协议、互换协议、互换期权，这些证券的价格都由一些特定期限的(Libor 或互换)利率决定．基于这个原因，我们事先设定以下时间点：

291

$$0 = T_0 < T_1 < T_2 < \cdots < T_M \tag{17.64}$$

且

$$\Delta_i = T_{i+1} - T_i, \quad i = 0,1,\cdots,M-1 \tag{17.65}$$

注意，由于一些计算天数的习惯问题，对 T_i 的设定将会使得每个 Δ_i 不尽相同．我们令 $P(t, T_n)$表示在 T_n 时刻到期的零息债券在 t 时刻的价格，其中 $T_n>t$，$n=1$，…，M. 类似地，我们用 $L_n(t)=L(t,T_n)$表示 t 时刻应用在$[T_n,T_{n+1}]$期间的远期利率，$n=1$，…，$M-1$. 根据方程(17.63)我们可以得到以下关于 Libor 利率的方程．

$$L_n(t)=\frac{P(t,T_n)-P(t,T_{n+1})}{\Delta_n P(t,T_{n+1})} \tag{17.66}$$

其中 $0\leqslant t\leqslant T_n$，$n=1$，…，$M-1$. T_i 时刻的 Libor 利率为

$$L_n(T_i)=\frac{P(T_i,T_n)-P(T_i,T_{n+1})}{\Delta_n P(T_i,T_{n+1})} \tag{17.67}$$

求解 $P(T_i,T_{n+1})$得

$$P(T_i,T_{n+1})=\frac{P(T_i,T_n)}{1+\Delta_n L_n(T_i)} \tag{17.68}$$

利用(17.68)我们可以写出 $P(T_i,T_n)$的表达式

$$P(T_i,T_n)=\frac{P(T_i,T_{n-1})}{1+\Delta_{n-1}L_{n-1}(T_i)} \tag{17.69}$$

将(17.69)带入(17.68)，可得

$$P(T_i,T_{n+1})=\frac{P(T_i,T_{n-1})}{(1+\Delta_{n-1}L_{n-1}(T_i))(1+\Delta_n L_n(T_i))} \tag{17.70}$$

重复该步骤，我们可以将 $P(T_i,T_{n+1})$与 Libor 利率的关系写成

$$P(T_i,T_{n+1})=\prod_{j=i}^{n}\frac{1}{1+\Delta_j L_j(T_i)} \tag{17.71}$$

从(17.71)我们可以发现，我们只能求债券在某些特定时间的价格．如果我们想求任意时间 t 时零息债券的价格 $P(t,T_N)$，$0\leqslant t\leqslant T_n$，该怎么办呢？定义指标 ℓ，使得 T_ℓ 是 t 时刻后的第一个时间节点，即满足 $T_{\ell-1}<t\leqslant T_\ell$. 易证

$$P(t,T_n)=P(t,T_\ell)=\prod_{j=\ell}^{n-1}\frac{1}{1+\Delta_j L_j(t)} \tag{17.72}$$

我们定义一个计价单位 B_t^*，叫做滚动银行账户，它构成方式如下：在 $T_0=0$ 时刻我们花 1 美元购买 $\frac{1}{P(0,T_1)}$ 份零息债券，T_1 时刻到期．到 T_1 时刻我们拥有的资金为 $\frac{1}{P(0,T_1)}$ 美元. 现在我们重复这一过程，购买 $\frac{1}{P(0,T_1)P(T_1,T_2)}$ 份零息债券，T_2 时刻到期；到 T_2 时刻我们拥有资金 $\frac{1}{P(0,T_1)P(T_1,T_2)}$ 美元．再购买 $\frac{1}{P(0,T_1)P(T_1,T_2)P(T_2,T_3)}$ 份零息债券，T_3 时刻到期. 一直重复这一过程，可以看到在 T_ℓ 时刻我们的资金为

$$\frac{1}{P(0,T_1)P(T_1,T_2)P(T_2,T_3)\cdots P(T_{\ell-1},T_\ell)} \tag{17.73}$$

根据(17.71)，我们可知上式等于

$$\prod_{j=0}^{\ell-1}(1+\Delta_j L_j(T_j)) \tag{17.74}$$

为了得出它在 t 时刻的价值，我们用 $P(t,T_\ell)$对其折现，得

$$B_t^*=P(t,T_\ell)\prod_{j=0}^{\ell-1}(1+\Delta_j L_j(T_j)) \tag{17.75}$$

把 B_t^* 作为计价单位，所得的测度就是即期测度 P^* . 将各可交易工具除以 B_t^* ，都会得到即期测度 P^* 下的鞅 . 我们用即期测度的计价单位除以零息债券价格，得到的结果记为 $D_n(t)$. 利用方程(17.72)和(17.75)，可以看到 $D_n(t)$满足

292

$$D_n(t)=\frac{P(t,T_\ell)\prod_{j=\ell}^{n-1}\frac{1}{1+\Delta_j L_j(t)}}{P(t,T_\ell)\prod_{j=0}^{\ell-1}(1+\Delta_j L_j(T_j))} \tag{17.76}$$

分子分母中都含有 $P(t,T_\ell)$，将它约掉，可得

$$D_n(t)=\frac{\prod_{j=\ell}^{n-1}\frac{1}{1+\Delta_j L_j(t)}}{\prod_{j=0}^{\ell-1}(1+\Delta_j L_j(T_j))} \tag{17.77}$$

在 Libor 市场模型中，为了在即期测度下得到 Libor 利率，我们建立以下一般的随机微分方程

$$\mathrm{d}L_n(t)=\mu_n(t)L_n(t)\mathrm{d}t+L_n(t)\sigma_n^{\mathrm{T}}(t)\mathrm{d}W_t,\ 0\leqslant t\leqslant T_n,\ n=1,\cdots,M \tag{17.78}$$

其中 $\mu_n(t)$是漂移率，是标量，而 $\sigma_n(t)$是波动率，是 d 维向量，它们都是适应过程 . W_t 是 d 维布朗运动，d 是因子的数量 . 我们知道在无套利条件下，$\mu_n(t)$可以写成 $\sigma_n(t)$的函数 . 现在我们的目标是要找到 $\mu_n(t)$的表达式 .

从(17.77)我们可以发现，如果我们对 $D_n(t)$取对数，可以得到更简单的对数形式表达式

$$\log D_n(t)=\log\left(\prod_{j=\ell}^{n-1}\frac{1}{1+\Delta_j L_j(t)}\right)-\log\left(\prod_{j=0}^{\ell-1}(1+\Delta_j L_j(T_j))\right) \tag{17.79}$$

$$=-\sum_{j=\ell}^{n-1}\log(1+\Delta_j L_j(t))-\sum_{j=0}^{\ell-1}\log(1+\Delta_j L_j(T_j)) \tag{17.80}$$

再对其求导，可以去掉第二项，得

$$\mathrm{d}\log D_n(t)=-\sum_{j=\ell}^{n-1}\mathrm{d}\log(1+\Delta_j L_j(t))+0 \tag{17.81}$$

应用伊藤引理，得

$$\mathrm{d}\log(1+\Delta_j L_j(t))=0+\frac{\Delta_j}{1+\Delta_j L_j(t)}\mathrm{d}L_j(t)-\frac{1}{2}\frac{\Delta_j^2}{(1+\Delta_j L_j(t))^2}(\mathrm{d}L_j(t))^2 \tag{17.82}$$

将(17.78)带入，可得

$$\mathrm{d}\log(1+\Delta_j L_j(t))=\frac{\Delta_j}{1+\Delta_j L_j(t)}(\mu_j(t)L_j(t)\mathrm{d}t+L_j(t)\sigma_j^{\mathrm{T}}(t)\mathrm{d}W_t) \tag{17.83}$$

$$-\frac{1}{2}\frac{\Delta_j^2}{(1+\Delta_j L_j(t))^2}L_j\sigma_j^{\mathrm{T}}(t)\sigma_j(t)\mathrm{d}t \tag{17.84}$$

合并同类项可得

$$\mathrm{dlog}D_n(t) = -\sum_{j=\ell}^{n-1}\mathrm{dlog}(1+\Delta_j L_j(t)) = \sum_{j=\ell}^{n-1}\left(\frac{\Delta_j^2 L_j^2(t)\sigma_j^{\mathrm{T}}(t)\sigma_j(t)}{2(1+\Delta_j L_j(t))^2} - \frac{\Delta_j \mu_j(t) L_j(t)}{1+\Delta_j L_j(t)}\right)\mathrm{d}t - \left(\sum_{j=\ell}^{n-1}\frac{\Delta_j L_j(t)\sigma_j^{\mathrm{T}}(t)}{1+\Delta_j L_j(t)}\right)\mathrm{d}W_t \tag{17.85}$$

$D_n(t)$是即期测度下的鞅，这意味着存在 $\mathbf{R}^d$ 上的过程 $v_n(t)$满足

$$\mathrm{d}D_n(t) = D_n(t)v_n^{\mathrm{T}}(t)\mathrm{d}W_t \tag{17.86}$$

对 $D_n(t)$使用伊藤引理，可以得到

$$\mathrm{dlog}D_n(t) = 0 + \frac{1}{D_n(t)}\mathrm{d}D_n(t) - \frac{1}{2}\frac{1}{D_n^2(t)}(\mathrm{d}D_n(t))^2 \tag{17.87}$$

293

将(17.86)带入，有

$$\mathrm{dlog}D_n(t) = v_n^{\mathrm{T}}(t)\mathrm{d}W_t - \frac{1}{2}v_n^{\mathrm{T}}(t)v_n(t)\mathrm{d}t = -\frac{1}{2}v_n^{\mathrm{T}}(t)v_n(t)\mathrm{d}t + v_n^{\mathrm{T}}(t)\mathrm{d}W_t \tag{17.88}$$

比较(17.88)和(17.85)的波动项，可得

$$v_n(t) = -\sum_{j=\ell}^{n-1}\frac{\Delta_j L_j(t)\sigma_j(t)}{1+\Delta_j L_j(t)} \tag{17.89}$$

现在我们要寻找漂移率 μ_j 的表达式．由于(17.88)和(17.85)相等，可得

$$\sum_{j=\ell}^{n-1}\left(\frac{\Delta_j^2 L_j^2(t)\sigma_j^{\mathrm{T}}(t)\sigma_j(t)}{2(1+\Delta_j L_j(t))^2} - \frac{\Delta_j \mu_j(t) L_j(t)}{1+\Delta_j L_j(t)}\right) = -\frac{1}{2}v_n^{\mathrm{T}}(t)v_n(t) \tag{17.90}$$

方便起见，我们定义

$$\eta_j(t) = \frac{\Delta_j L_j(t)}{1+\Delta_j L_j(t)} \tag{17.91}$$

利用(17.89)可得

$$\frac{1}{2}\sum_{j=\ell}^{n-1}\eta_j^2(t)\sigma_j^{\mathrm{T}}(t)\sigma_j(t) - \sum_{j=\ell}^{n-1}\eta_j(t)\mu_j(t) = -\frac{1}{2}\left(\sum_{j=\ell}^{n-1}\eta_j(t)\sigma_j^{\mathrm{T}}(t)\right)\left(\sum_{j=\ell}^{n-1}\eta_j(t)\sigma_j(t)\right) \tag{17.92}$$

令 $n=\ell+1$，则

$$\frac{1}{2}\eta_\ell^2(t)\sigma_\ell^{\mathrm{T}}(t)\sigma_\ell(t) - \eta_\ell(t)\mu_\ell(t) = -\frac{1}{2}\eta_\ell^2(t)\sigma_\ell^{\mathrm{T}}(t)\sigma_\ell(t) \tag{17.93}$$

$$\eta_\ell(t)\mu_\ell(t) = \frac{1}{2}\eta_\ell^2(t)\sigma_\ell^{\mathrm{T}}(t)\sigma_\ell(t) = \frac{1}{2}\eta_\ell^2(t)\sigma_\ell^{\mathrm{T}}(t)\sigma_\ell(t) \tag{17.94}$$

$$= \eta_\ell^2(t)\sigma_\ell^{\mathrm{T}}(t)\sigma_\ell(t) \tag{17.95}$$

因此

$$\mu_\ell(t) = \eta_\ell(t)\sigma_\ell^{\mathrm{T}}(t)\sigma_\ell(t) \tag{17.96}$$

令 $n=\ell+2$，则

$$\eta_\ell(t)\mu_\ell(t) + \eta_{\ell+1}(t)\mu_{\ell+1}(t) = \frac{1}{2}\eta_\ell^2(t)\sigma_\ell^{\mathrm{T}}(t)\sigma_\ell(t) + \frac{1}{2}\eta_{\ell+1}^2(t)\sigma_{\ell+1}^{\mathrm{T}}(t)\sigma_{\ell+1}(t)$$

$$+\frac{1}{2}(\eta_\ell(t)\sigma_\ell^{\mathrm{T}}(t)+\eta_{\ell+1}(t)\sigma_{\ell+1}^{\mathrm{T}}(t))((\eta_\ell(t)\sigma_\ell(t)+\eta_{\ell+1}(t)\sigma_{\ell+1}(t)) \tag{17.97}$$

展开后利用(17.96)重新组合，可得

$$\eta_{\ell+1}(t)\mu_{\ell+1}(t)=\eta_{\ell+1}^2(t)\sigma_{\ell+1}^{\mathrm{T}}(t)\sigma_{\ell+1}(t)+\eta_{\ell+1}(t)\eta_{\ell+1}(t)\eta_\ell(t)\sigma_{\ell+1}^{\mathrm{T}}(t)\sigma_\ell(t) \tag{17.98}$$

所以

$$\mu_{\ell+1}(t)=\eta_{\ell+1}(t)\sigma_{\ell+1}^{\mathrm{T}}(t)\sigma_{\ell+1}(t)+\eta_\ell(t)\sigma_{\ell+1}^{\mathrm{T}}(t)\sigma_\ell(t) \tag{17.99}$$

$$=\sigma_{\ell+1}^{\mathrm{T}}(t)(\eta_\ell(t)\sigma_\ell(t)+\eta_{\ell+1}(t)\sigma_{\ell+1}(t)) \tag{17.100}$$

$$=\sigma_{\ell+1}^{\mathrm{T}}(t)\sum_{j=\ell}^{\ell+1}\eta_j(t)\sigma_j(t) \tag{17.101}$$

$$=\sigma_{\ell+1}^{\mathrm{T}}(t)v_{\ell+2}(t) \tag{17.102}$$

通过归纳法，我们可以证明

$$\begin{aligned}\mu_n(t)&=\sigma_n^{\mathrm{T}}(t)\sum_{j=\ell}^{n}\eta_j(t)\sigma_j(t)=\sigma_n^{\mathrm{T}}(t)v_{n+1}(t)\\&=\sigma_n^{\mathrm{T}}(t)\sum_{j=\ell}^{n}\frac{\Delta_jL_j(t)\sigma_j(t)}{1+\Delta_jL_j(t)}=\sum_{j=\ell}^{n}\frac{\Delta_jL_j(t)\sigma_n^{\mathrm{T}}(t)\sigma_j(t)}{1+\Delta_jL_j(t)}\end{aligned} \tag{17.103}$$

294

所以我们可以得到 P^* 无套利远期 Libor 利率的动态变化公式

$$\mathrm{d}Ln(t)=\Big(\sum_{j=\ell}^{n}\frac{\Delta_jL_j(t)\sigma_n^{\mathrm{T}}(t)\sigma_j(t)}{1+\Delta_jL_j(t)}\Big)L_n(t)\mathrm{d}t+L_n(t)\sigma_n^{\mathrm{T}}(t)\mathrm{d}W_t,$$
$$0\leqslant t\leqslant T_n,\quad n=1,\cdots,M \tag{17.104}$$

17.3.3 一些含义

在上一节，我们选取了一个在 $t=t_3$ 到期的债券来得到综合概率(测度)，而在这个新测度下的鞅等式对定价更为方便．选择 $B_{t_1}^S$ 作为标准化因子的部分原因是为了使过程简便．

实际上，任何其他资产也都可以作为标准化变量．但是，由于 $B_{t_1}^S$ 在 t_3 时刻到期，它在 t_3 时刻的价值为常数．由于这一简单的原因，我们能很容易地得到 π 测度下的条件期望值．由于该性质，依赖于 u，d 的项 $L_{t_2}^{ij}$ 和 $C_{t_2}^{ij}$ 关于 t_1 时刻的信息集为常数，如方程(17.28)包含的信息．由于它们是常数，这些系数可以被提取到期望算子外面．这是一个重要结果，它使我们不必考虑即期利率和受未来利率影响的价格的相关性．而且，由于这一点，我们不会遇到随机折现因子．

选择这种标准化方法还有一个重要原因．选择适当的标准化之后，远期利率如 F_{t_1} 将成为鞅，并且是 L_{t_2} 等即期利率未来值的无偏估计．这意味着我们能够用相应的远期利率来代替即期利率的未来值，从而找到各种受利率影响的证券的当前无套利价格．

在本节的最后，我们将本章的内容应用到两个连续时间下的定价例子上．

1. FRA 合约

假定我们现有一份 FRA 合约，合约规定在某个未来日期 $T+\Delta$ 支付 $(F_t-L_T)N$，

$t<T$. 其中 N 为名义金额，F_t 为随机变量 L_T 的远期价格．F_t 是根据合约签订时刻 t 的观察确定的.

现金流在未来时刻 $T+\Delta$ 支付，其现值 V_t 可以由常用的鞅等式给出，其中未来现金流用无风险利率 r_s 折现，$t\leqslant s\leqslant T+\delta$. 在风险中性测度下我们有：

$$V_t = E^Q[\mathrm{e}^{-\int_t^{t+\delta} r_u \mathrm{d}u}(F_t - L_T)N\delta] \tag{17.105}$$

现在，我们已知远期合约启动时不涉及任何现金交换．那么在合约启动时，[1]我们有

$$V_t = 0 \tag{17.106}$$

在什么条件下该价格为 0 呢？我们通过选择适当的 F_t 来使右边的期望消失．如果认为即期利率是确定的，这就很简单．只要将折现因子提出来，就能够得到 F_t 的值

$$V_t = \mathrm{e}^{-\int_t^{t+\delta} r_u \mathrm{d}u} E^Q[F_t - L_T]N\delta \tag{17.107}$$

然后令 V_t 等于 0，可得：

$$0 = E^Q[F_t - L_T] \tag{17.108}$$

若 F_t 能使远期合约的现值为 0，则必须满足

$$F_t = E^Q[L_T] \tag{17.109}$$

295

也就是说，当即期利率为确定值，远期价格等于在风险中性测度下对 L_T 的“最佳”预测．在 Black-Scholes 框架下，当对股票期权定价时假设利率为常数．但是这样的假设在对利率敏感型证券定价时不可行．最重要的原因是这些证券的作用就是对冲利率风险．显然，我们假设即期利率是确定性的，也不是非常合适．但是如果不假设 r_t 是确定性的，折现因子就不能提到外面，也就不能在测度 Q 下使用方程(17.59).

而远期测度能够提供一种非常方便的解决方法．利用折现债券 $B(t,T+\delta)$的无套利价格，我们能够写出在远期测度下的定价方程：

$$V_t = E^{\pi}[B(t,T+\delta)(F_t - L_T)N\delta] \tag{17.110}$$

其中 $\delta>0$. 这里 $B(t, T)$表示 t 时刻的价值；所以它能够提到期望算子外面：

$$V_t = B(t,T+\delta)E^{\pi}[(F_t - L_T)]N\delta \tag{17.111}$$

现在，利用 $V_t=0$，得：

$$F_t = E^{\pi}[L_T] \tag{17.112}$$

使用该方程，我们就能很方便地求 F_t 的值．关键的一点是，在求该平均时，我们使用的是远期测度 π 而不是风险中性概率 Q.

2. 上限协议

为了体现远期测度的优势，我们的第二个例子是上限协议．现有关于 Libor 利率 L_t 的上限协议，设 C_t 是该上限协议的现价，δ 为期限，K 为合约中确定的上限利率．假定标的金额为 $N=1$ 且协议在 T 时刻到期．我们令 $\delta=1$，L_t 上的下标记号除外．

根据以上信息，该上限协议的购买者在 T 时刻将获得的收入为

---

[1] 任何可能被要求的利润不属于现金交换，而是作为未来交易实施的保证．

$$C_T = \max[L_{T-\delta} - K, 0]$$

之前我们说过，这种金融工具能帮购买者将利率限制在 K 以下．$0<\delta<1$，且右边的值与 δ 成正比．

我们应该怎样对该上限协议定价呢？假设我们要使用风险中性概率 Q. 我们知道无套利价格为：

$$C_t = E^Q[e^{-\int_t^T r_u du} \max[L_{T-\delta} - K, 0]] \tag{17.113}$$

我们还知道，在 $T-\delta$ 时刻 $F_{T-\delta}$ 与 $L_{T-\delta}$ 相等，[㊀] 所以我们考虑用 F_t 作为“标的”. 毕竟随着时间推移，该变量最终会与未来即期利率 $L_{T-\delta}$ 趋于一致．该过程叫做远期 Libor 过程．

这意味着我们可以对远期利率对数正态形式的动态变化建模，并且使用在原始概率 P 下的维纳过程 W_t，

$$dF_t = \mu F_t t + \sigma F_t dW_t$$

这和 Black-Scholes 的设定一样，下面我们用 Black-Scholes 中的方法来确定 C_t.

按这种方法，我们的第一步是选择风险中性概率 Q. 但是这会造成一个问题，F_t 不是在 Q 下的鞅．所以，由于我们在 Q 下选择概率和维纳过程 W_t^*，远期利率的微分方程为：

$$dF_t = \mu^* F_t t + \sigma F_t dW_t^*$$

其中 μ^* 是满足 Girsanov 定理的新的风险调整漂移率．在 Q 下，漂移率不是直接给出的．在 Black-Scholes 情况下，标的股票价格的漂移率用已知的(常数)即期利率 r 替代．而这里的情况不同，我们需要确定一个未知的漂移率，而这非常困难．

296

现在我们考虑如果使用远期测度 π，那远期利率的微分方程将会发生什么变化．在用 $B(t, T)$ 标准化得到的远期测度下，$T-\delta$ 的远期利率 F_t 将成为鞅．[㊁] 因此我们有：

$$dF_t = \sigma F_t dW_t^\pi$$

其中 dW_t^π 是测度 π 下的维纳过程．该 SDE 具有非常实用的性质，其漂移率等于零，且 F_t 是 $L_{T-\delta}$ 的无偏估计：

$$F_t = E^\pi[L_{T-\delta}]$$

这样就不用确定一个未知的漂移率了．我们能按照 Black-Scholes 的方法直接对该上限协议定价．[㊂]

3. 作为工具的标准化

上面我们着重从利率敏感型证券定价的角度，讨论了标准化方法和测度选择的重要含义. 这在金融衍生品的数学方面有什么含义呢?

从前面的讨论中我们可以看到，基础变量实际上是状态价格 $\{\psi^{ij}\}$. 如果不存在套利机

㊀ 任何时候，期限为 δ 的短期贷款的远期利率等于该时段的即期利率．

㊁ 我们必须认识到，若使用别的标准化方法，这个远期利率可能不再是鞅．

㊂ 还有一个区别在于这里所用的单位．如果使用远期利率的微分方程，我们不必将该上限协议的收益折现到当前. 这与 Black-Scholes 环境不同，在后者中，股票价格的 dS_t 需要被折现到 t 时刻．

会，状态价格唯一存在，且大于零．一旦确定了这一点，金融分析师就有很大的灵活性来考虑他可选的鞅测度．综合概率既可以取经典风险中性测度 Q，也可以取远期测度 π，这取决于我们研究的是哪种金融工具．所以，选取哪种测度成了分析师的又一个问题．

实际上，根据 Girsanov 定理，我们可以根据定价问题的需要选择各种概率．考虑我们刚才使用 B_t^S 得到的标准化方法以及对应的测度 π. 显然，我们还可以用期限较长的债券 B_t 来进行标准化，并得到新的概率 $\widetilde{\pi}$：

$$\widetilde{\pi}^{ij} = \frac{1}{B_t}\psi^{ij} \tag{17.114}$$

一旦使用 B_t 进行标准化，所有在 T 时刻到期的工具价格都将成为鞅．

注意比率

$$\frac{\pi^{ij}}{\widetilde{\pi}^{ij}} = \left[\frac{\dfrac{1}{B_t^s}}{\dfrac{1}{B_t}}\right] \tag{17.115}$$

可以写为

$$\pi^{ij} = \widetilde{\pi}^{ij}\left[\frac{\dfrac{1}{B_t^s}}{\dfrac{1}{B_t}}\right] \tag{17.116}$$

这样我们就能从一个测度转换到另一个测度．

这种调整在利率敏感型工具定价中对我们来说有用吗？答案是肯定的．当我们研究的金融工具依赖于一个以上的 L_T 和不同的期限 T 时，我们可以从一种远期测度入手，再关于另一种测度求导数，然后我们就可以引入合适的“修正项”.

17.4 小结

本章我们介绍了标准化和远期测度的概念．这些工具在衍生证券定价中扮演着重要角色，带来了极大便利．它们不仅是重要的理论概念，还是现实市场中的资产定价常用工具. 对于涉及未来执行、未来现金流的衍生品，这些工具尤为实用．

297

下面我们列出了一些主要结果．当我们利用无违约风险的折现债券 $B(t,T)$得到远期测度 π_T 时，将有以下 3 个结果：

- 在用零息债券的无套利价格 $B(t,T)$进行标准化之后，系统内所有资产的价格都将成为 π_T 下的鞅．
- 无需额外的标准化，期限与 $B(t,T)$相同的远期价格本身也是鞅．[⊖]
- 关于到期日 T 的折现因子将成为确定性的量，可以从定价方程中提取出来．

⊖ 在这种情况下，远期价格 F_{t_1} 本身也是相对于未来时间而言的，不像 C_t，是 t 时刻的价格．

17.5 参考阅读

对于背景较强的读者，Musiela 和 Rutkowski(1997)是非常好的著作．虽然它的内容相比本章对数学的要求更高，但它能使读者获益匪浅．此外 Pliska(1997)的最后一章也值得推荐．Pliska 在离散时间下介绍了这些概念，这一点和我们是一样的．

17.6 习题

1. 假设以下关于即期利率 r_t 的信息：

- r_t 满足：

$$dr_t = \mu r_t dt + \sigma r_t dW_t$$

- 年漂移率为

$$\mu = 0.01$$

- 年波动率为

$$\sigma = 12\%$$

- 假设当前即期利率为 6%.

(a) 假设要对为期一年的金融工具定价．确定合适的时间间隔 Δ，使二叉树有 5 步．

(b) 在这种情况下，μ 和 d 具有怎样的含义？

(c) 画出即期利率的二叉树．

(d) 在二叉树中，向上、向下移动的概率分别是多少？

2. 假设在 $t=0$ 时刻有 4 份无违约风险的零息债券 $P(t,T)$，期限分别为 1 到 4 年：

$$P(0,1) = 0.94, \quad P(0,2) = 0.92, \quad P(0,3) = 0.87, \quad P(0,4) = 0.80$$

(a) 能否找到一个符合以上价格的即期利率树？请讨论．

(b) 画一个期限结构与上述信息一致的树．

(c) 问题(a)和(b)中的求树法有什么区别？如果有的话．

3. 用 Mathematica、Maple 或者 Matlab 软件取 10 个标准正态随机数．假设利率服从以下 SDE：

$$dr_t = 0.02 r_t dt + 0.06 r_t dW_t$$

当前即期利率假设为 6%.

298

(a) 将上述 SDE 离散化．

(b) 计算以下期望的估计值，取时间间隔 $\Delta=0.04$，并利用你选的随机数，在真实概率下求期望．

$$E\left[e^{-\int_0^1 r_S dS} \max(r_1 - 0.06, 0) \right]$$

(c) 计算 $E\left[e^{-\int_0^1 r_S dS} \right]$ 的样本均值，再用它乘以 $E[\max(r_1 - 0.06, 0)]$ 的样本均值．我们会得到什么结果？

(d) 哪种方法是正确的？

(e) 我们能在计算债券价格时使用该结果吗？

(f) 特别地，我们怎样再能知道上述 SDE 中利率的动态变化是无套利的呢？

(g) 如果我们将测度变为风险中性测度 Q，上面的利率方程会发生什么变化？

(h) 假设你已知一系列债券的无套利价格，怎样利用这些信息得到无套利 r_t 的动态变化？

4. 我们观察到 1 年期的零息债券现价为 0.97 美元(在 $t=1$ 时支付 1 美元). 我们还观察到当前的远期利率是 2.5%. 在远期测度下，对关于 6 个月 Libor 的期限为 6 个月的上限协议定价. 假设在远期测度下 F_t 的动态变化为：$dF_t=\sigma F_t dW_t$. 假设波动率为 9%，执行利率为 3%.

5. 根据 Vasicek 模型，对 1 年期的零息债券定价

$$dr_t = \alpha(\mu - r_t)dt + \sigma dW_t \tag{17.117}$$

其中 $r_0=0.01$，$\alpha=0.2$，$\mu=0.01$，$\sigma=0.05$.

6. 证明：互换协议的价格等于上限协议的价格减去下线协议的价格.

7. 证明：互换期权支付方的价格减去收入方的价格等于远期互换协议的价格.

8. 考虑一个在未来生效的互换协议. 协议在 T_n 生效，在 T_N 到期. 假设每个计息期的长度为 δ. 在互换协议中，支付是延迟的，故第一笔支付发生在 $T_{n+1}=T_n+\delta$.

(a) 假设名义本金为 1 美元，证明该互换协议在 $t\leqslant T_n$ 的总价值为

$$\Pi_p(t) = P(t,T_n) - P(t,T_N) - S\delta\sum_{j=n+1}^{N}P(t,T_j)$$

其中 S 是合约规定的固定(年)利率.

(b) 利用(a)中的结论，求远期互换利率 $R_{\text{swap}}(t)$的表达式，$t\leqslant T$. 其中 $R_{\text{swap}}(t)$在利率 S 基础上有 $\Pi_p(t)=0$.

(c) 证明：远期互换利率 $R_{\text{swap}}(t)$可以写成简单远期利率的加权平均，即

$$R_{\text{swap}}(t) = \sum_{j=n+1}^{N}\omega_j(t)F(t;T_{j-1},T_j)$$

并求出 $\omega_j(t)$. $F(t;T_{j-1},T_j)$是 t 时刻$[T_{j-1},T_j]$期间简单远期利率的值.

9. 远期平价互换利率 $y_{n,N}(t)$的定义为

$$y_{n,N}(t) = \frac{P(t,T_n)-P(t,T_N)}{\sum_{j=n+1}^{N}\delta P(t,T_j)} = \frac{P(t,T_n)-P(t,T_N)}{P_{n+1,N}(t)}$$

$P_{n+1,N}(t)$叫做基点价值(PVBP).

299

互换期权的持有者有参与某个互换合约的权利而非义务. 若互换期权的期权到期日为 T_n，互换到期日为 T_N，则即为 $T_n\times T_N$－互换期权. 该互换期权的总时间是 T_n+T_N. 由于互换期权的持有者拥有参与某个互换合约的权利而非义务，可以将它看作关于互换利率的看涨期权. 该期权在期权到期日 T_n 的收入为

$$\begin{aligned}[V_{n,N}^{\text{Payer}}(T_n)]^+ &= \left[\{1-P(T_n,T_N)\} - \kappa\sum_{j=n+1}^{N}\delta P(T_n,T_j)\right]^+ \\ &= [y_{n,N}(T_n)P_{n+1,N}(T_n) - \kappa P_{n+1,N}(T_n)]^+ = P_{n+1,N}(T_n)[y_{n,N}(T_n)-\kappa]^+\end{aligned}$$

其中 κ 是互换期权的执行利率，第二步直接来自远期互换利率的定义. 我们令 $B_t = \exp\left(\int_0^t r_s ds\right)$ 为 t 时刻的货币市场账户. 假设无套利，互换期权在 t 时刻的价值记作 PS_t，可以表示为以下风险中性条件期望

$$\frac{PS_t}{B_t}E^Q\left\{\frac{[V_n^{\text{Payer}}(T_n)]^+}{B_{T_n}}\right\}$$

$$\frac{PS_t}{B_t} = E^Q\left\{\frac{P_{n+1,N}(T_n)}{B_{T_n}}[y_{n,N}(T_n)-K]^+\right\}$$

(a) 利用 $P_{n+1,N}(t)$作为计价单位，得到新的概率测度 $P^{n+1,N}$，叫做互换测度．

(b) 在互换测度下证明

$$PS_t = P_{n+1,N}(t)E_t^{P^{n+1,N}}\{[y_{n,N}(T_n) - K]^+\}$$

注意在互换测度下，对应的互换利率 $y_{n,N}(t)$是鞅．通过改变计价单位，我们可以清晰地看出为什么互换期权可以看作关于互换利率的期权．

10. 用归纳法证明，在无套利假设下，漂移率 μ_1，…，μ_n 必须满足方程(17.103)．证明：$D_{n+1}(t)$是鞅当且仅当 $L_n(t)D_{n+1}(t)$是鞅，并利用伊藤引理得到(17.103)．

300

第 18 章 期限结构建模及相关概念

18.1 引言

上一章非常重要，它对利率敏感型证券的金融基本定理进行了讨论，得到了一些新结果. 我们利用简单的模型讨论了标准化的问题、Libor 工具基础上远期测度的使用、不同期限债券同时存在时的处理方法等. 在具体使用第 17 章得出的结果前，我们需要先回过头来对一些基本概念进行更加详细的讨论.

首先我们要做两件事. 固定收益方面的新概念相较标准 Black-Scholes 领域的简单概念要更零碎、直观性更差. 因此需要先定义这些固定收益概念，再来推导定价公式. 否则我们将难以理解债券定价公式的推导过程.

接下来我们还需要引入一些重要的套利关系，它们在利率敏感型证券的定价中反复出现. 下一章会介绍对利率敏感型证券定价的两种方法，分别是*经典方法*和 *Heath-Jarrow-Morton 方法*. 我们的目的是说明这些不同方法背后的原理，并且突出讨论它们的相似点和不同点. 为此，我们需要先引入一些存在于即期利率、债券价格和远期利率之间的新的套利关系.

我们要了解的第一种套利关系是短期存款账户投资与债券投资之间的关系. 假设它们都无违约风险. 长期债券价格与短期存款账户中的存款的连续滚动投资之间存在什么关系呢?[㊀] 显然，当我们购买期限较长的债券时，所获得的收益就不仅仅是隔夜或隔月收益了. 在这段较长的时间里，可能发生很多风险事件，而它们可能会反向影响债券价格. 而隔夜投资几乎不受风险事件的影响，因为投资者的钱在第二个交易日就会被返还，可以按更高的隔夜利率进行再投资. 为使风险厌恶的投资者持有长期债券，相对隔夜投资来说，似乎长期债券应该支付一定的溢价. 在 Black-Scholes 环境中，变换到风险中性测度可以消除风险溢价，并给出定价公式. 对于利率敏感型证券和随机即期利率也可以这样做吗? 我们将会看到这个问题的答案是肯定的. 事实上，对利率敏感型证券定价的经典方法利用的就是这种特殊的套利关系.

第二种套利关系是固定收益类产品所特有的. 固定收益市场中有很多几乎完全相同的流动性工具，它们只有期限不同，比如一系列仅期限不同的折价债券. 类似的，也存在期限不同的远期利率. 可以证明利率敏感型工具的这种多维特性使得零息债券与远期利率之间存在着复杂的套利关系. 事实上，对于 k 维的债券价格向量，我们可以利用套利讨论将它与远期利率向量联系在一起. 这些套利关系形成了 Heath-Jarrow-Morton 利率敏感型证券定价方法的基础.

㊀ 实践中，期限最短的投资只获得隔夜利息.

因此无论使用哪种方法，本章的内容都是讨论利率敏感型证券定价必备的背景知识．

18.2 主要概念

我们从一些概念开始讨论，其中有些在前面介绍过．用 $B(t,T)$表示 T 时刻到期的折价债券在 $t<T$ 时的价格．r_t 仍表示无风险借贷的瞬时即期利率．这里所谓的瞬时即期利率是指贷款是发生在 t 时刻、并且在无穷小时间段 $\mathrm{d}t$ 后被归还的．这里即期利率也是无风险的，因为不存在违约风险，该瞬时投资的收益率是确定的[⊖]．这两个概念我们已经在前面见到过．

302

我们现在要定义的第一个新概念是折价债券 $B(t,T)$的连续复利收益率 $R(t,T)$．已知债券现价 $B(t,T)$，其面值为 1 美元，$R(t,T)$定义如下：

$$B(t,T)=\mathrm{e}^{-R(T,t)(T-t)} \tag{18.1}$$

$R(t,T)$是"现在投资 $B(t,T)$美元，$[T-t]$后获得 1 美元"对应的收益率．这里使用指数函数表示连续复利．注意，这种债券价格与收益率之间的关系一一对应的．已知其中一个，我们就可以求出另一个，它们也用相同的下标 T 和 t 表示．

下面我们定义连续复利远期利率 $F(t,T,U)$．它表示 T 时刻开始、$U>T$ 时刻到期的贷款利率．尽管现金交易发生在未来的 T 时刻和 U 时刻，该利率是在 t 时刻由合约确定的．该利率是连续复利率说明实际发生的利息计算需要使用指数函数．事实上，如果在 T 时刻贷款 1 美元，U 时刻归还的钱就应为：

$$\mathrm{e}^{F(t,T,U)(U-T)} \tag{18.2}$$

注意，$F(t,T,U)$有 3 个时间指标，而每个折价债券都只有两个时间指标．这说明我们需要使用两种期限分别为 T 和 U 的债券 $B(t,T)$和 $B(t,U)$，才能得到远期利率与债券价格之间的关系．这两种债券合起来具有时间指标(t,T,U)．

18.2.1 3 条曲线

前一节定义的基本概念可用于定义被市场参与者广泛使用的 3 条曲线．这 3 条曲线分别是收益率曲线、贴现曲线和信用价差曲线．由于超出了本书所考虑的范围，我们在这里省略了固定收益市场中应用最广泛的互换曲线，不对这种工具进行考虑．我们仅讨论这些曲线涉及的数学原理．远期曲线包含一系列期限各异的远期贷款利率，我们将在下一章中讨论．

1. 收益率曲线

收益率曲线由收益率 $R(t,T)$与折价债券价格 $B(t,T)$之间的关系确定．我们有：

$$B(t,T)=\mathrm{e}^{-R(T,t)(T-t)},\quad t<T \tag{18.3}$$

其中 $B(t,T)$是 T 时刻到期折价债券的无套利价格．因此要得到债券收益率 $R(t,T)$，我们需要先知道到它的价格．然后我们就可以通过(18.3)计算连续复利收益率：

⊖ 但是同前面一样，即期利率本身可以是随机变量，如果进行滚动投资，该投资也可能存在市场风险．

$$R(T,t)=\frac{\log(1)-\log B(t,T)}{T-t}=\frac{-\log B(t,T)}{T-t} \tag{18.4}$$

只要 $t<T$，我们就有 $0<B(t,T)<1$. 因此 $\log[B(t,T)]$小于零，进而，$R(t,T)$一定大于零．

假设在 t 时刻存在期限 $T\in[t,T_{\max}]$的一系列零息债券，其中 $T_{\max}$ 表示市场上的最长期限．用集合$\{B(t,T),\ T\in[t,T_{\max}]\}$表示这些债券的价格．对于该集合中任意 $B(t,T)$，我们都可以利用(18.4)得到相应收益率 $R(t,T)$. 于是有下面的定义．

定义 21 $\{B(t,T),\ T\in[t,T_{\max}]\}$的收益率形成的曲线称为收益率曲线．

收益率曲线将特定风险等级的债券收益率与相应期限联系起来．

上面给出的收益率曲线定义对从业者使用的收益率概念进行了延伸．所观察到的收益率曲线给出了与有限个不同期限对应的(比如)国债收益率．假设时间是连续的，且在任意时刻 t 都存在纯折价债券连续统．只要 $T<T_{\max}$，投资者总可以买入或卖出 T 时刻到期的流动债券．这些期限的取值范围可以从 t 时刻的下一个瞬间，

303

$$T=t+\mathrm{d}t \tag{18.5}$$

到可能的最长期限 $T=T_{\max}$，从而形成了一条连续的收益率曲线．根据该假设，对于任意 $T<T_{\max}$，不需对相应收益率进行“插值”，因为它们可以在市场上直接观察得到．

2. 贴现曲线

尽管“收益率曲线”这一术语很流行，市场上应用得更多的是贴现曲线．

定义 22 无违约风险的零息债券价格$\{B(t,T),\ T\in[t,T_{\max}]\}$及属于同一风险等级的期限连续统，形成了贴现曲线．

对于一般的现金流计算而言，使用贴现曲线更方便．事实上，用$\{cf_{T_1},\cdots,cf_{T_n}\}$表示任意 $T_1<T_2<\cdots<T_n$ 时刻得到的现金流，将这些在 T_i 时刻收到的数额分别乘以对应的 $B(t,T_i)$，就可以得到该现金流的现值 CF_t. 利用期限分别为 T_i 的无套利零息债券价格可以轻松地计算出贴现值．该现值等于

$$CF_i=\sum_{i=1}^{n}B(t,T_i)cf_{T_i} \tag{18.6}$$

这背后的原因非常简单，$B(t,T_i)$就是在 T_i 时刻得到的 1 美元的无套利现值．该贴现因子可以在市场中找到．因此，贴现曲线在市场从业者的日常工作中起着关键作用．

3. 信用价差曲线

显然，对于给定风险等级的债券，收益率曲线和贴现曲线是有效的．考察$\{B(t,T),\ T\in[t,T_{\max}]\}$时，我们假设这些债券的违约风险是相同的，否则收益率之间的差异将与期限以外的因素有关．

因此，对于每个风险等级而言，我们都可以得到一条不同的收益率(贴现)曲线．这些收益率(贴现)曲线之间的差异就是信用价差，即相同期限下，信用风险更大的债券需要支付的溢价. 代表不同风险等级的不同收益率曲线的共存性就得出了所谓的信用价差曲线．

定义 23 已知两条收益率曲线，一条为无违约风险债券的收益率曲线$\{R(t,T),$

$T\in[t,T_{\max}]\}$，另一条为违约概率已知的债券收益率曲线$\{\widetilde{R}_t^T, T\in[t,T_{\max}]\}$，称其溢价$\{s(t,T)=\widetilde{R}_t^T-R(t,T),\ T\in[t,T_{\max}]\}$为信用价差曲线．

一些从业者确实更喜欢使用信用价差与期限之间的联系，而不是讨论收益率曲线本身．如果交易工具以价差而不是基础利率为标的，那么使用信用价差曲线将更实用．我们在本书中省略了对信用工具的讨论，并假设不存在违约风险．因此只有一种风险等级，并且违约风险为 0.

18.2.2 收益率曲线的运动

在讨论更实质的问题之前，我们想要先比较一下收益率曲线本身的移动与沿着某条特定收益率曲线的移动．已知收益率曲线$\{R(t,T),T\in[t,T_{\max}]\}$在 T 时刻连续，我们可以在 t 时刻讨论两种不同的增量变化．首先，在任意瞬间 t，我们感兴趣的是当 T 变化很小的量 $\mathrm{d}T$ 时，特定的 $R(t,T)$发生的变化．这里我们将特定债券的期限从 T 调整到 $T+\mathrm{d}T$. 也就是说，我们在沿着同一条收益率曲线移动．据此，如果收益率曲线是连续且光滑的，我们可以得到导数：

304

$$\frac{\mathrm{d}R(t,T)}{\mathrm{d}T}=g(T) \tag{18.7}$$

这就是收益率曲线$\{R(t,T),T\in[t,T_{\max}]\}$的斜率．这些量如图 18-1 所示．$g(T)$是在 T 时刻与连续收益率曲线相切直线的斜率．图 18-2 给出的是相应的贴现曲线．

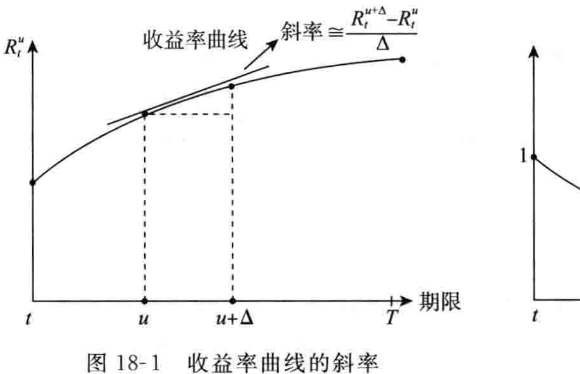


图 18-1 收益率曲线的斜率

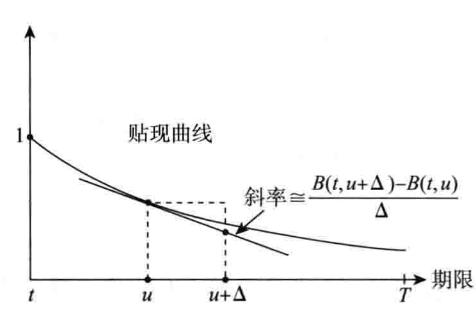


图 18-2 贴现曲线的斜率

收益率曲线通常按斜率小于零、斜率大于零和斜率等于零进行分类，它们也会呈现出"驼峰"形．当曲线形状变化时，斜率也发生变化．有一点很重要，就是要意识到 T 时刻的增量变化不会涉及未知的随机波动．如果 t 时刻每个 $R(t,T)$都是已知的，那么这就是一个 t 时不同期限债券的价格所构成的曲线．同样，由于这些运动中不含有维纳增量，该导数可以按标准法则处理，不需借助伊藤引理．

我们考虑的第二种增量变化是时间参数 t 变化时所发生的变化．时间 t 发生变化导致的收益率 $R(t,T)$的增量变化会包含随机波动．当 t 发生变化时，时间随之推移，产生新的

维纳增量．而这样的随机波动会影响即期利率，因而，收益率曲线会发生移动．当 t 变化了 $\mathrm{d}t$ 后，通常整条收益率曲线都将发生变化，理解这一点很重要．因此，固定收益工具的动态变化本质上是曲线的动态变化，而不是单个随机过程的动态变化．隐含的套利限制将比 Black-Scholes 环境中更加复杂，毕竟我们要确保新的曲线下不存在套利机会．

还要注意的是 $\mathrm{d}R(t,T)$ 这类随机微分，由于含有维纳成分，需要使用伊藤引理进行计算．现在我们已经为介绍基本的定价公式(即债券定价公式)做好了准备，该公式的使用将贯穿本书的后半部分．

18.3 债券定价公式

我们在本节开始讨论本章的第一个实质性问题，逐步推导出 T 时刻到期的无违约风险零息债券 $B(t,T)$ 的无套利定价公式．先从一个简化后的例子(其瞬时即期利率是常数)开始讨论，然后讨论随机无风险利率．这样可以使套利关系的讨论更容易理解．

18.3.1 常数即期利率

令即期利率 r_t 为常数：

$$r_t = r \tag{18.8}$$

则 T 时刻支付 1 美元的无违约风险纯折价债券价格为：

$$B(t,T) = \mathrm{e}^{-r(T-t)} \tag{18.9}$$

这里 r 是连续复利瞬时利率，该函数是 t 时刻的贴现因子．当 $t=T$ 时，该指数函数等于 1，与该债券的到期价值相等．当 $t<T$ 时，该指数函数都小于 1. 因此等式(18.9)右边代表 T 时刻的 1 美元按连续复利常数利率 r 贴现后的现值． 305

对于这些投资工具，投资者可以有以下选择．他(或她)现在可以对无风险存款账户进行投资，使其在 T 时刻价值为 1 美元．也可以购买 T 时刻到期的折价债券，现在支付 $B(t,T)$ 美元，在 T 时刻也获得 1 美元．显然这两种工具都没有违约风险，T 时刻得到的收益相同．两种工具都不存在中间支付．如果利率是常数，且不存在违约风险，那么任意承诺在 T 时刻支付 1 美元的债券价格都必须等于初始的无风险借款额 $\mathrm{e}_t^{-r(T-t)}$ 美元，即必须有：

$$B(t,T) = \mathrm{e}^{-r(T-t)} \tag{18.10}$$

否则将存在套利机会．对该环境中的套利讨论进行一下回顾将很有帮助． 306

我们先假设 $B(t,T)>\mathrm{e}^{-r(T-t)}$．那么我们可以在 $[t,T]$ 期间卖空债券，并利用所得收益投资无风险借贷．T 时刻，(债券的)空头头寸价值为 1 美元，而无风险借贷可以得到 1 美元．因此 T 时刻的净现金流将等于 0. 但投资者在 t 时刻时仍有一定的结余，因为

$$B(t,T) > \mathrm{e}^{-r(T-t)}$$

另一种可能的情形是 $B(t,T)<\mathrm{e}^{-r(T-t)}$．那么我们将在 t 时刻借入美元，并以价格 $B(t,T)$ 买入债券．到期时净现金流仍等于 0. 债券支付的 1 美元可以用于还贷．但在 t 时刻时会有净收益：

$$e^{-r(T-t)} - B(t,T) > 0$$

消除这些套利机会的唯一方法就是债券等价公式满足：

$$B(t,T) = e^{-r(T-t)} \tag{18.11}$$

因此这种关系并不是一种定义或假设，而是债券价格与储蓄账户间无套利机会所施加的限制条件．注意，我们在得到该等式时没有用到金融基本定理，即使使用，得到的结果仍然相同．[㊀]

18.3.2 随机即期利率

如果瞬时即期利率 r_t 是随机利率，(18.9)中的定价公式将发生变化．假设 r_t 为无穷小区间$[t,t+dt]$的无风险利率，则 r_t 在 t 时刻是已知的，但其未来值将随时间波动．这里可以应用金融基本定理得到 $B(t,T)$与随机即期利率 r_t，$t\in[t,T]$之间的套利关系．

利用第 17 章中介绍的方法．当前债券价格 $B(t,T)$用储蓄账户上 1 美元的现值进行标准化，接下来考虑债券到期时的价值(1 美元)，用储蓄账户的价值来对其进行标准化．该值等于 $e^{\int_t^T r_s ds}$，它在 T 时刻等于 1 美元，而在整个 $s\in[t,T]$期间都是按瞬时利率 r_s 滚动计息的．用 1 美元除以储蓄账户中的这个值，得到 $e^{-\int_t^T r_s ds}$.

根据第 17 章，该标准化后的债券价格在风险中性测度 Q 下必须是鞅．因此有

$$B(t,T) = E^Q\left[e^{-\int_t^T r_s ds}\right] \tag{18.12}$$

其中 $e^{-\int_t^T r_s ds}$ 项也表示 1 美元面值的贴现因子．

接下来将对该公式进行进一步讨论．首先，(18.12)给出的债券定价公式具有其他重要含义．债券价格依赖于未来所有短期利率 r_s，$t<s<T$. 也就是说我们可以这样看：t 时刻的收益率曲线包含所有与未来短期利率相关的信息．[㊁]

其次，使用哪种概率测度计算这些数学期望也是个问题．也许有人认为国债等级的债券是无风险资产，不存在需要消除的风险溢价，因此不需要使用等价鞅测度．通常来说这种观点是错的．因为当利率是随机时，国债价格将包含市场风险．它们依赖于即期利率的未来变化，而这种变化是随机的．要消除与此相关的风险溢价需要使用等价鞅测度来计算(18.12)中的表达式．

307

我们现在使用 $0<\Delta$ 的离散区间来讨论该公式，并与连续时间情况联系起来，更好地解释债券定价公式的原理．

离散时间

考虑离散时间框架下的一个特殊情形：三期债券．如果 Δ 表示短于一年的时间区间，t 表示现在，那么三期折价债券的价格为：

㊀ 见本章练习题 2.

㊁ 当已知信息集 I_t 时，条件期望给出了均方误差意义下的最佳预测．

$$B(t,t+3\Delta) = E^Q\left[\frac{1}{(1+r_t\Delta)(1+r_{t+\Delta}\Delta)(1+r_{t+2\Delta}\Delta)}\right] \tag{18.13}$$

其中 r_t 为已知的当前贷款即期利率，该贷款起始于 t 时刻、结束于 $t+\Delta$ 时刻，而 $r_{t+\Delta}$，$r_{t+2\Delta}$ 为未来两期的未知即期利率．不同于连续时间情况，它们都是单利利率，按市场惯例乘以 Δ.

根据(18.13)，债券价格等于到期收益的贴现值．贴现因子是随机的，需要进行(条件)期望运算．该期望是在风险中性概率测度 Q 下计算的．用无风险借贷对债券在 t 时刻和 $t+3\Delta$ 时刻的值进行标准化．与前面一样，对于 t 时刻，我们用债券价格除以 1；对于 $t+3\Delta$ 时刻，我们用到期时债券价格(1 美元)除以按未来即期利率 $r_{t+\Delta}$，$r_{t+2\Delta}$滚动投资的收益．该条件期望的条件是 t 时刻已获得的信息集，而该信息集包含 r_t 的当前值．

18.3.3 连续时间

我们逐渐将当前的设定变为连续时间的情况．当我们把三期变为 n 期时，公式(18.13)变为：

$$B(t,t+3\Delta) = E^Q\left[\frac{1}{(1+r_t\Delta)(1+r_{t+\Delta}\Delta)\cdots(1+r_{t+n\Delta}\Delta)}\right] \tag{18.14}$$

其中 Δ 满足 $T=t+n\Delta$. 当 r_j 很小时，存在近似式：

$$\frac{1}{1+r_j\Delta} \approx e^{-r_j\Delta} \tag{18.15}$$

接下来分别将它应用到(18.14)右边每个比率中，得到近似式

$$\frac{1}{(1+r_t\Delta)(1+r_{t+\Delta}\Delta)\cdots(1+r_{t+n\Delta}\Delta)}$$

$$= e^{-r_t\Delta}e^{-r_{t+\Delta}\Delta}\cdots e^{-r_{t+n\Delta}\Delta} \tag{18.16}$$

$$= e^{-(r_t+r_{t+\Delta}+\cdots+r_{t+n\Delta})\Delta} \tag{18.17}$$

当 $\Delta\to 0$ 时，

$$e^{-\sum_{i=1}^{n}(r_{t+\Delta})\Delta} \to e^{-\int_t^T r_s \mathrm{d}s} \tag{18.18}$$

因此当 $\Delta\to 0$ 时，我们从离散时间贴现变为利用即期利率变量对连续时间贴现．于是离散时间贴现因子变为指数函数．由于利率是连续变化的，指数中需要使用积分．这样就得到了连续时间债券定价公式：

$$B(t,T) = E^Q[e^{-\int_t^T r_s \mathrm{d}s}]$$

18.3.4 收益率与即期利率

我们也可以推导出收益率 $R(t,T)$与短期利率 r_t 之间的关系．利用(18.12)和(18.3)将未来短期利率与 t 时刻收益率曲线联系起来．有： 308

$$e^{-R(t,T)(T-t)} = E^Q[e^{-\int_t^T r_s \mathrm{d}s}] \tag{18.19}$$

取对数：

$$R(t,T) = \frac{-\log E^Q[e^{-\int_t^T r_s \mathrm{d}s}]}{T-t} \tag{18.20}$$

我们看到债券收益率可以(大致地)被债券存续期内平均即期利率所刻画．事实上，在常数即期利率的特例中，

$$r_s = r, \quad t \leqslant s \leqslant T \tag{18.21}$$

我们得到

$$R(t,T) = \frac{-\log E^Q\left[\mathrm{e}^{-\int_t^T r_s \mathrm{d}s}\right]}{T-t} = \frac{-\log \mathrm{e}^{-r(T-t)}}{T-t} \tag{18.22}$$

$$= r \tag{18.23}$$

因此，只要即期利率是常数，收益率就等于即期利率．

18.4 远期利率与债券价格

我们在本节得到另一种套利关系，该关系说明了远期利率与债券价格之间的关系．我们会看到这种关系在现代固定收益理论中起着关键作用．

$F(t,T,U)$表示t时刻订立的、起始于T时刻且结束于$U>T$时刻的远期利率．

与前面一样，要推导出$B(t,T)$与$F(t,T,U)$之间的关系，我们还需要在U时刻到期的第二种债券$B(t,U)$. $F(t,T,U)$表示包含t时刻关于未来T与U时刻之间信息的市场价格．我们认为债券$B(t,T)$是包含到T时刻为止全部相关信息．另一方面，期限较长的债券$B(t,U)$是包含到U时刻为止全部信息的价格．因此，理论上我们应该可以从$B(t,U)$和$B(t,T)$中得到全部关于$F(t,T,U)$的必要信息．和前面一样，我们先利用区间Δ得到在离散时间下的这种关系，然后取极限得到连续时间的情况．

18.4.1 离散时间

我们先从两期债券开始讨论．如果Δ表示很小但非无穷小的时间区间，且t表示现在，那么两期债券的价格将为$B(t,t+2\Delta)$. 该债券将在$t+2\Delta$时刻到期时产生1美元现金流．于是我们在t时刻支付$B(t,t+2\Delta)$购买该债券，这笔投资将在$t+2\Delta$时刻产生1美元收益．

假设存在远期贷款的流动市场，考虑t时刻的替代投资．进行远期贷款，贷款开始于$t+\Delta$时刻，在$t+2\Delta$时刻支付的利息为$F(t,t+\Delta,t+2\Delta)\Delta$. 假设贷款数额可以使得我们在$t+2\Delta$时刻得到1美元，因此$t+\Delta$时刻贷款本金记为$B^*_{t+\Delta}$:[⊖]

$$B^*_{t+\Delta} = \frac{1}{1+F(t,t+\Delta,t+2\Delta)\Delta} \tag{18.24}$$

这是$t+\Delta$时刻的数目．我们需要利用当前的即期利率将它贴现到t时刻．远期贷款在t时刻的价值B^*_t为：

309

$$B^*_t = \frac{1}{1+r_t\Delta}\left[\frac{1}{1+F(t,t+\Delta,t+2\Delta)\Delta}\right] \tag{18.25}$$

⊖ 读者可能会注意到下面的表达式中远期利率$F(t,T,T+\Delta)$都乘以了Δ. 这一点是必要的，因为$F(\cdot)$是年利率，而Δ是很短的随机区间．按市场惯例，Δ期间得到的远期利息不应该为$F(t,T,T+\Delta)$，而应该是$F(t,T,T+\Delta)$乘以Δ. 例如，如果远期年利率是6%，而一年包含360天，那么三月期贷款得到的收益将为$6\times1/4\%$.

最后，由于 t 时刻利率 r_t 也是平凡定义的远期利率 $F(t, t, t+\Delta)\Delta$，则 B_t^* 等于：[⊖]

$$B_t^* = \frac{1}{[1+F(t,t,t+\Delta)\Delta]}\frac{1}{[1+F(t,t+\Delta,t+2\Delta)\Delta]} \tag{18.26}$$

据此，如果我们在 t 时刻按利率 r_t 投资一定数额，在 $t+\Delta$ 时刻按事先约定的利率 $F(t,t+\Delta,t+2\Delta)$滚动这笔投资，在 $t+2\Delta$ 时将获得 1 美元收益．而这种策略的收益与购买债券 $B(t,t+2\Delta)$完全相同．因此，如果两种策略中的信用风险相同，一定有

$$B(t,t+2\Delta) = B_t^*$$

或

$$B(t,t+2\Delta) = \frac{1}{[1+F(t,t,t+\Delta)\Delta]}\frac{1}{[1+F(t,t+\Delta,t+2\Delta)\Delta]} \tag{18.27}$$

注意，由于右边所有量在 t 时刻都是已知的，因此不需要在该公式中使用任何期望运算．债券价格与当前的远期利率之间的关系是确定的．

当该套利关系不成立时会发生什么呢？我们可以卖空贵的投资、买入便宜的投资．两种头寸的支付与收益将在 $t+2\Delta$ 时相互抵消，但在 t 时刻获得利润．因此会存在套利机会．

18.4.2 连续时间

假设我们考虑离散的 n 期，每期长度均为 Δ，则 $T=t+n\Delta$. 该公式变为

$$B(t,T) = \frac{1}{[1+F(t,t,t+\Delta)\Delta]\cdots[1+F(t,t+(n-1\Delta),t+n\Delta)\Delta]} \tag{18.28}$$

当 $F(t,T,U)$和 Δ 都很小时，使用近似式，有

$$\frac{1}{[1+F(t,T,U)\Delta]} \approx \mathrm{e}^{-F(t,T,U)\Delta} \tag{18.29}$$

于是 $B(t,T)$可以写成

$$B(t,T) \approx [\mathrm{e}^{-F(t,t,t+\Delta)\Delta}][\mathrm{e}^{-F(t,t+\Delta,t+2\Delta)\Delta}]\cdots[\mathrm{e}^{-F(t,t+(n-1)\Delta,t+n\Delta)\Delta}] \tag{18.30}$$

而指数项的积可以简化，于是

$$B(t,T) \approx \mathrm{e}^{-F(t,t,t+\Delta)\Delta-F(t,t+\Delta,t+2\Delta)\Delta-\cdots-F(t,t+(n-1)\Delta,t+n\Delta)\Delta} \tag{18.31}$$

$$= \mathrm{e}^{-\sum_{i=1}^{n}F(t,t+(i-1)\Delta,t+i\Delta)\Delta} \tag{18.32}$$

也就是说我们可以令 $\Delta\to 0$，并增加区间数目，以得到连续时间下瞬时远期利率与债券价格之间的关系，

$$B(t,T) = \mathrm{e}^{-\int_t^T F(t,s)\mathrm{d}s} \tag{18.33}$$

已知循环条件都已满足．现在 $F(t,s)$是 t 时刻订立、s 时刻开始、无穷小时间区间 $\mathrm{d}s$ 后结束的瞬时远期利率．因此当 $\Delta\to 0$ 时，我们从离散时间变为连续时间贴现．于是得到了两个结果．首先，离散贴现因子被指数函数所取代．其次，我们要使用的是瞬间远期利率而不是离散远期利率．由于瞬时远期利率可能不同，因此需要在指数项中包含积分．这里仍然没有期望运算，这是因为 $F(t,s)$在 t 时刻都是已知量．

310

⊖ 任意现在就开始的远期贷款都被称为平凡远期贷款．

公式为：

$$B(t,T) = e^{-\int_t^T F(t,s)ds} \tag{18.34}$$

以瞬时远期利率函数的形式给出了无违约风险零息债券的价格．

相反我们也可以将 $F(t,T,U)$ 写成债券价格的函数．当期限为 T 且 $U=T+\Delta$ 时，我们更倾向于这样写．[⊖] 考虑 $B(t,T)$ 和 $B(t,T+\Delta)$ 两种债券，它们的期限只相差很小的时间区间 $\Delta>0$. 再次写出公式(18.32)：

$$B(t,T) = e^{-\int_t^T F(t,s)ds} \tag{18.35}$$

及

$$B(t,T+\Delta) = e^{-\int_t^{T+\Delta} F(t,s)ds} \tag{18.36}$$

对两个等式取对数，然后相减：

$$\log B(t,T) - \log B(t,T+\Delta) = -\int_t^T F(t,s)ds + \int_t^{T+\Delta} F(t,s)ds \tag{18.37}$$

$$= \int_T^{T+\Delta} F(t,s)ds \tag{18.38}$$

假设 Δ 很小，使得我们可以认为 $F(t,T)$ 在很小的时间区间 $[T,T+\Delta]$ 内是常数．则有：

$$\log B(t,T) - \log B(t,T+\Delta) \approx F(t,T)\Delta \tag{18.39}$$

取极限后，该等式就变为：

$$F(t,T) = \lim_{\Delta\to 0}\frac{\log B(t,T) - \log B(t,T+\Delta)}{\Delta} \tag{18.40}$$

即瞬时远期利率 $F(t,T)$ 与贴现曲线取对数后的导数密切相关．

通过类似的讨论，我们可以得到类似的连续复利远期利率 $F(t,T,U)$：[⊜]

$$F(t,T,U) = \frac{\log B(t,T) - \log B(t,U)}{U-T}$$

其中 $F(t,T,U)$ 是起始于 $T<U$ 时刻、结束于 U 时刻的连续复利远期利率，该合约订立于 t 时刻．

显然，令 $T\to U$ 可以得到瞬时远期利率 $F(t,T)$：

$$F(t,T) = \lim_{T\to U} F(t,T,U) \tag{18.41}$$

从这些讨论中显然可以看出 $F(t,T)$ 的存在是建立在贴现曲线（即债券价格连续统）关于 T（到期日）可微的基础上的．利用(18.39)，并假设满足其他技术性条件，我们有

$$F(t,t) = r \tag{18.42}$$

即起始于当前 t 时刻的贷款的瞬时远期利率等于即期利率 r_t.

18.5 小结

311

我们对目前为止所得到的结论进行回顾．我们基本导出了债券价格 $B(t,T)$、债券收益

⊖ 这样有助于理解后面 Heath-Jarrow-Morton(HJM)套利条件的推导．

⊜ 本章前面讨论离散时间情况时使用 $F(t,T,U)$ 表示单利远期利率．当变为连续时间情况并引入指数函数时，相同的符号表示连续复利率．对这两种概念使用不同的符号表示更加准确．但由于本章的记法已很复杂，故我们使用了相同的符号进行表示．

率 $R(t,T)$ 及远期利率 $F(t,T,U)$ 与即期利率 r_t 之间的 3 种关系．

第一种关系由定义得出．已知债券价格，定义连续复利收益率 $R(t,T)$ 为：

$$R(t,T)=\frac{-\log B(t,T)}{T-t} \tag{18.43}$$

第二种关系是对债券价格应用本书前半部分原理所得到的结果，即在风险中性测度 Q 下金融衍生品的收益应当等于这些工具当前的无套利价格，但要用瞬时利率 r_t 来进行贴现．债券 $B(t,T)$ 到期时支付 1 美元，其贴现值为

$$\left[e^{-\int_t^T r_s \mathrm{d}s}\right]$$

即期利率 r_t 是随机的，在风险中性测度下对它进行(条件)期望运算，得到：

$$B(t,T)=E^Q\left[e^{-\int_t^T r_s \mathrm{d}s}\right] \tag{18.44}$$

因此第二种关系是建立在无套利条件的基础上的定价公式．即给定一个合适的 r_t 模型后，就可以用它来计算债券 $B(t,T)$ 的证券市场价格．

第三种关系由前一节导出．再次使用套利讨论，我们看到债券 $B(t,T)$、$B(t,U)$ 的(无套利)价格与连续复利远期利率 $F(t,T,U)$ 之间密切相关：

$$F(t,T,U)=\frac{\log B(t,T)-\log B(t,U)}{U-T},\quad t<T<U \tag{18.45}$$

也可以把它当做定价等式来使用，但如果给定的是 $F(t,T,U)$，我们将得到含有两个未知量 $B(t,T)$、$B(t,U)$ 的等式．因此，在我们将它作为定价公式前，需要至少先知道 $B(t,T)$、$B(t,U)$ 中 1 个值．增加其他远期利率并没有什么帮助，因为每个远期利率等式都含有一个未知债券价格．[⊖]

总之，第一种关系就是一个简单的定义关系，不能用于定价．但另外两种是建立在套利原则的基础上的，并在流动的、功能完备的市场上成立．它们形成了两种对利率敏感工具定价方法的基础．所谓的*经典方法*利用的是第二种关系，而新近流行的 Heath-Jarrow-Morton(HJM)*方法*使用的是第三种关系．我们将在下一章中学习这些内容．

18.6 参考阅读

关于离散时间固定收益模型，读者们可以参考 Jarrow(1996). Rebonato(1998)中对利率模型进行了非常全面的回顾．另外一本很好的参考资料是 Risk(1996). 对于近期的讨论可以参考 Jegadeesh 和 Tuckman(2000).

18.7 习题

1. 考虑关于即期利率 r_t 的 SDE

$$\mathrm{d}r_t=\alpha(m-r_t)\mathrm{d}t+\sigma \mathrm{d}W_t \tag{18.47}$$

⊖ 假设引入另一个含有 $B(t,U)$ 的等式：

$$F(t,U,S)=\frac{\log B(t,U)-\log B(t,S)}{S-U},\quad t<U<S \tag{18.46}$$

我们就有两个方程，但有三个未知数：$B(t,T)$、$B(t,U)$ 和 $B(t,S)$. 我们要需要知道一些其他信息．

假设参数 α，μ，σ 已知，W_t 仍为维纳过程．

(a) 证明

312

$$E[r_s \mid r_t] = \mu + (r_t - \mu)e^{\alpha(s-t)}, \quad t < s \tag{18.48}$$

$$V[r_s \mid r_t] = \frac{\sigma^2}{2\alpha}(1 - e^{2\alpha(s-t)}), \quad t < s \tag{18.49}$$

(b) 当 $s \to \infty$ 时，这两个等式的条件均值及方差有什么含义？

(c) 假设利率风险的市场价格为常数 λ(即 Girsanov 变换利用 $\sigma\lambda$ 调整漂移项)．使用给出的债券价格函数，证明在 s 时刻到期的债券漂移项和扩散项分别为

$$\mu^B = r_t + \frac{\sigma\lambda}{\alpha}(1 - e^{\alpha(s-t)}) \tag{18.50}$$

$$\sigma^B = \frac{\lambda}{\alpha}(1 - e^{\alpha(s-t)}) \tag{18.51}$$

(d) 到期时债券波动率会发生什么变化？它是可预测的吗？

(e) 到期时漂移系数会发生什么变化？它是可预测的吗？

(f) 最后，期限极长($s \to \infty$)的债券漂移系数和扩散系数有什么含义？

2. 考虑两阶段情况，且在 $t=0$，1，2 时，都存在两种可能状态．一种是以无风险利率 $r_i(i=0, 1)$进行无风险借贷，另一种是按现价 B_0 购买两阶段债券，该债券在 $t=2$(到期)时支付 1 美元．

(a) 构建一个 2×4 的系统，其中状态价格为 $\psi^{ij}(i, j=u, d)$，使得我们通过该系统可以得出储蓄账户和债券 B 的无套利价格．

(b) 证明可以在该条件下得到风险中性概率测度．

(c) 证明：若将储蓄账户正规化，则债券的无套利价格将为

$$B = E_0^Q\left[\frac{1}{(1+r_0)(1+r_1)}\right]$$

3. 零息债券的期望值是多少？

$$B(0,T) = E(e^{-\int_t^s r_u d_u}) \tag{18.52}$$

根据下面的动态原理：

$$dr_t = \alpha(\mu - r_t)dt + \sigma dW_t \tag{18.53}$$

该零息债券将服从什么分布？

提示：考虑 $\int_t^s r_u du \approx \sum_t^s r_u du$ 的分布．除此以外，$\mathrm{var}\left(\int_t^s r_u du\right) = \frac{\sigma^2}{\alpha^2}\left[s - t + \frac{1 - e^{-2\alpha(s-t)}}{2\alpha} - 2\frac{1 - e^{-\alpha(s-t)}}{\alpha}\right]$.

4. 考虑利率衍生品，若短期利率大于 0.01，则它在 T 时刻到期时支付 1 美元，否则支付 0. 当 $T=1$ 年时，利用 Vasicek 模型

$$dr_t = \alpha(\mu - r_t)dt + \sigma dW_t \tag{18.54}$$

313

对该衍生品进行定价．其中 $r_0=0.01$，$\alpha=0.2$，$\mu=0.01$，$\sigma=0.05$.

第 19 章　固定收益产品的经典定价法和 HJM 定价法

19.1　引言

市场上对利率敏感型证券定价的方法有两种，它们分别基于前一章所介绍的两种套利关系．事实上，第 18 章详细讨论了债券定价公式

$$B(t,T) = E^{Q}[\mathrm{e}^{-\int_t^T r_s \mathrm{d}s}]$$

该公式给出了无违约风险的折价债券 $B(t,T)$在风险中性测度 Q 下的无套利价格．这种即期利率 r_t 与债券价格 $B(t,T)$之间的关系仅在套利机会不存在时成立．

第 18 章中的第二种套利关系是瞬时远期利率 $F(t,T)$与债券价格之间的关系：

$$B(t,T) = \mathrm{e}^{-\int_t^T F(t,T)\mathrm{d}s}$$

显然这两种关系都可以用于计算利率敏感型证券的无套利价格．

市场实践通常从一系列(可以被认为是无套利的)债券价格$\{B(t,T)\}$出发，利用上面两种关系中的任意一种，倒推得到 r_t 的模型或一系列$\{F(t,s),s\in[t,T]\}$的模型．由于上面两种关系的前提都是无套利条件，因此得到的 r_t 模型或瞬时远期利率都将是"对风险进行调整后"的结果，即它们在风险中性测度 Q 下都是有效的．

所谓的经典方法使用的是第一种套利关系，从$\{B(t,T)\}$中得到即期利率 r_t 风险调整后的模型．这种方法将涉及对即期利率模型中漂移系数的建模和波动率的调整．这里通常假设 r_t 具有马尔可夫性质．

而 Heath-Jarrow-Morton(HJM)方法使用的是第二种套利关系，得到 k 维瞬时远期利率$F(t, T)$的无套利模型．该方法不需要对漂移系数建模，但需要调整波动率．它是更为一般的、但在实践中应用较少的方法．HJM 方法不需要对即期利率建模，但是它却能显示即期利率 r_t 往往并不具有马尔可夫性．

我们在本章中讨论的是从业者所使用的利率敏感型证券定价方法．由于篇幅所限，省略了数值问题和定价计算的细节．感兴趣的读者可以参考一些这方面的优秀著作．我们的重点在于理解这两种完全不同的方法背后的思想．

19.2　经典方法

债券价格与瞬时即期利率之间的关系为

$$B(t,T) = E^{Q}[\mathrm{e}^{-\int_t^T r_s \mathrm{d}s}] \tag{19.1}$$

市场从业人员可以按(至少)两种方法使用这一关系．

第一种方法，如果存在准确的无套利贴现曲线$\{B(t,T)\}$，我们就可以利用(19.1)倒推得到即期利率 r_t 的无套利模型．然后利用该即期利率模型的无套利性质对除债券外的其他利率衍生品进行定价．

我们也可以用另一种方法．如果不存在任何关于贴现曲线$\{B(t,T)\}$的可靠数据，我们可以先利用利率的历史数据进行估计，提出一个合适的即期利率 r_t 的无套利模型，然后利用(19.1)得到非流动债券和其他利率敏感型衍生品的公平市场价格．这两种方法都被称为利率衍生品定价的*经典方法*．我们将会看到，无论使用两种方法中的哪一种，经典方法的基础都是对瞬时利率 r_t 建模，第一种方法从可靠的债券价格$\{B(t,T)\}$出发，而第二种方法从 r_t 过程本身可得到的数据出发．

这两种方法都不够直接，因此我们先来看一些简单的例子．

19.2.1 例 1

先考虑直接对 r_t 建模的情况．

假设在一个经济体中折价债券的交易不活跃，我们可以认为 r_s 等于常数 r. 即

$$r_s = r, \quad s \geqslant t$$

则我们可以将(19.1)写成：

316

$$B(t,T) = E^Q\left[e^{-\int_t^T r ds}\right]$$

由于 r 是常数，我们取平凡期望，得到：

$$B(t,T) = e^{-r(T-t)}$$

因此，从 r_t 的假设模型出发，我们得到了债券定价公式，即依赖于已知量 T，t 和 r 的闭形公式．

我们可以利用该等式对非流动债券进行定价．例如，假设 $r=0.05$. 则我们可以得到 1、2、3、4 年期折价债券的价格：

$$B(t,t+1) = 0.95$$
$$B(t,t+2) = 0.90$$
$$B(t,t+3) = 0.86$$
$$B(t,t+4) = 0.82$$

如果我们对 r_t 是常数的假定是正确的，这些债券价格将都是无套利的．

然而要注意的是，如果我们提出的是 r_t 的非确定性模型，使用相同步骤将会产生问题. 事实上这种方法要求知道即期利率过程在风险中性测度 Q 下的漂移系数．在上面的例子中，r_t 是常数，因此它在测度 Q 下漂移系数为 0.

19.2.2 例 2

假设我们不知道随机过程 r_t 在现实中属于哪种类型，即我们的目的是根据市场上交易的流动债券观测值确定该过程．特别地，假设我们观测到下面的贴现曲线：

$$B(t,t+1)=0.95$$
$$B(t,t+2)=0.90$$
$$B(t,t+3)=0.86$$
$$B(t,t+4)=0.82$$

则我们可以从这些价格推断出 r_t 服从 SDE

$$\mathrm{d}r_t=a(r_t,t)\mathrm{d}t+b(r_t,t)\mathrm{d}W_t$$

其中

$$a(r_t,t)=0,\quad b(r_t,t)=0$$

即 r_t 过程实际上就是常数 r.㊀

利用该信息，我们可以对标的为 r_t 或 $B(t,T)$ 的利率敏感型衍生品进行定价．例如，由于 r_t 为常数，其相应债券期权的无套利价格等于 0.

19.2.3　一般情形

假设我们有准确的无套利贴现曲线 $\{B(t,\ u),\ t<u\leqslant T\}$．则(19.1)说明相同的即期利率过程 r_s 必须满足下面的等式：㊁

$$B(t,T_0)=E^Q\left[\mathrm{e}^{-\int_t^{T_0}r_s\mathrm{d}s}\right] \tag{19.2}$$

$$B(t,T_1)=E^Q\left[\mathrm{e}^{-\int_t^{T_1}r_s\mathrm{d}s}\right] \tag{19.3}$$

$$\cdots\cdots \tag{19.4}$$

$$B(t,T_n)=E^Q\left[\mathrm{e}^{-\int_t^{T_n}r_s\mathrm{d}s}\right] \tag{19.5}$$

其中

$$T_0<T_1<\cdots<T_n \tag{19.6}$$

是 $n+1$ 种期限，已知这 $n+1$ 种期限对应债券的准确无套利价格．

我们来对这些方程进行一些讨论．已知定义期望运算 $E^Q[\cdot]$ 的鞅测度，尽管 T_i 不同，这些等式的右边都依赖于相同的即期利率过程 r_t. 市场会确定这 $n+1$ 个等式左边的值．从业人员需要为利率过程 r_t 确定一个模型，使该模型同时满足所有这些等式．如何确保所选的 r_t 模型在这 $n+1$ 个等式中保持一致性呢？这确实不是一项简单的任务．我们来对这之中的难点进行一下说明． 317

考虑如下要求．首先，我们假设即期利率模型为：

$$\mathrm{d}r_t=a(r_t,t)\mathrm{d}t+b(r_t,t)\mathrm{d}V_t \tag{19.7}$$

其次，我们需要选择 $a(r_t,t)$，$b(r_t,t)$ 及驱动过程 V_t 的概率行为，满足(19.2)～(19.5)等

㊀ 这些价格与例 1 中的价格完全相同．

㊁ 我们假设债券面值为 1.

式系统.[1]考虑两个例子.

1. **几何布朗运动 SDE**

要看出(19.2)—(19.5)等式系统中的隐含复杂性，我们假设即期利率 r_t 满足 Q 测度下维纳过程驱动的几何布朗运动 SDE：

$$dr_t = \mu r_t dt + \sigma r_t dW_t \tag{19.8}$$

于是我们可以假设(19.7)中漂移系数和扩散系数为：

$$a(r_t,t) = \mu r_t, \sigma(r_t,t) = \sigma r_t, V_t = W_t \tag{19.9}$$

这将会产生一个问题.当 $t \to \infty$ 时，服从该模型的即期利率过程将趋于正无穷或负无穷(取决于 μ 的符号).r_t 的百分比波动率也将为常数.显然，这并不是隔夜利率模型的理想性质.首先，利率没有趋势.其次，现实中利率的波动率是即期利率 r_t 的复杂非线性函数，而不是常数.[2]

但是让我们先把这两点困难放在一边，考虑上面提到的问题：即如何选择 μ 和 σ 使得(19.2)～(19.5)系统中的等式被同时满足？

这不是一项简单的工作.事实上，如果给定债券价格合理准确的观测值$\{B(t,T_i)，i=0,\cdots,n\}$，我们就得到了(19.2)～(19.5)中 $n+1$ 个等式的左边项.但是在利率模型中我们可选择的自由参数只有 μ 和 σ，因此我们需要通过确定两个未知量来满足含有 $n+1$ 个等式的系统.而这一点只有在债券价格$\{B(t,T_i),i=0,\cdots,n\}$之间存在很强的相互关系(使得这些等式中有 $n-1$ 个冗余)时才是有可能做到的，那样该系统将会变为两个等式，且只含有两个未知数 μ 和 σ 存在满足无套利贴现曲线$\{B(t,T)\}$.

但是这种假设我们可以在流动市场上观察到 $n+1$ 个债券价格间存在强相互关系的处理方式是否具有吸引力呢？显然(19.8)假设的随机利率过程在实际定价中是不充分的，我们还需要找到其他模型.

2. **均值回归模型**

几何布朗运动 SDE 可能不适合描述即期利率的动态原理，但我们可以从上面的讨论中有所启发.首先要为 r_t 选择合适的 SDE，接下来确定(调整)该模型的参数使得该即期利率模型满足流动市场中已知的贴现曲线$\{B(t,T_i)\}$.如果满足这一点，且观察到的折价债券价格$\{B(t,T_i),i=0,\cdots,n\}$是无套利的，那么得到的即期利率过程 r_t 模型也应该是无套利的.它可以用于对利率敏感型衍生品进行定价.

318

因此我们这里关心的是是否可以给出比前面的几何布朗运动更符合现实的 SDE.考虑带有均值变量 θ_t 和平方根扩散项的均值回归即期利率过程：

$$dr_t = \lambda(\theta_t - r_t)dt + \sigma\sqrt{r_t}\,dW_t \tag{19.10}$$

假设参数 θ_t 在不同时段 t 可以等于不同的已知值.这样的假设增加了(19.2)—(19.5)系统

---

[1] 注意这里的隐含假设.即期利率的增量仅依赖于当前的 r_t，因此即期利率具有马尔可夫性质.我们下面将会看到，这是满足套利条件的固定收益市场中的一种特例.

[2] 我们令 V_t 为维纳过程.这一点可能也是不合适的，因为现实世界中即期利率过程可能含有跳跃.

中自由参数的数目．例如，在离散环境下：

$$t_0 < t_1 < \cdots < t_m \tag{19.11}$$

利率模型中存在 $m+3$ 个可以选择的自由参数，即

$$\{\theta_{t_0}, \theta_{t_0}, \cdots, \theta_{t_m}, \lambda, \sigma\} \tag{19.12}$$

这使得将利率过程与已知贴现曲线$\{B(t, T_i), i=0, \cdots, n\}$相适应时更具有灵活性．[㊀]事实上，我们不仅可以将 r_t 过程应用于债券价格，还可以将它应用于债券波动率．[㊁]具体例子见 Hull 与 White(1990).

除此以外，不同于几何布朗运动过程，均值回归过程在合适条件下是确定的，不会当 $t\to\infty$ 时变得无穷大．同样，已知无穷小变化，如果采用这里的扩散系数，由均值回归模型得到的 r_t 过程不会小于零．

19.2.4　即期利率模型的使用

假设我们根据定价公式：

$$B(t,T) = E^Q[\mathrm{e}^{-\int_t^T r_s \mathrm{d}s}]$$

确定了即期利率 r_t 的无套利模型，该如何使用该模型呢？

我们在本章之初简要提到过这个问题的答案．债券定价公式用于从已知的期限结构中得到即期利率的无套利模型，这样做的原因在于我们可以用得到的模型对其他利率敏感型证券直接进行定价，而无需查看这些证券在对应市场的条件下得到它们的无套利价格．[㊂]

为具体理解即期利率模型的用途，考虑如下框架．已知期限结构 $B(t,T)$，利用它得出 r_t 的无套利模型：

$$\mathrm{d}r_t = \tilde{a}(r_t, t)\mathrm{d}t + b(r_t, t)\mathrm{d}W_t$$

其中漂移系数 $\tilde{a}(r_t, t)$上方具有波浪符，因为我们假设它是对利率风险进行调整后的漂移系数．W_t 是风险中性测度 Q 下的维纳过程．考虑两种情况．

1. 单因子模型

假设我们要对仅受 r_t 影响的衍生工具定价．将它的价格记为 $C(r_t, t)$，到期日为 T，到期收益由已知函数 $G(r_T, T)$确定：

$$C(r_T, T) = G(r_T, T)$$

可以直接使用定价公式：

$$C(r_t, t) = E^Q[\mathrm{e}^{-\int_t^T r_s \mathrm{d}s} G(r_T, T)]$$

我们可以使用蒙特卡洛方法计算该期望，也可以解出其封闭形式的解（如果存在），或将其　319

㊀ 参数 n 与 m 不一定相同．

㊁ 即调整(19.10)中 SDE 的自由参数，使得通过(19.1)得到的 $B(t,T)$的波动率与这些债券在流动期权市场上的波动率相符．

㊂ 我们至少可以说出 3 种不同的用途，但由于篇幅所限，这里不提及其他用途：(1)使用可交易债券 $B(t,T)$复制可交易工具 $C(t,T)$．(2)$C(t,T)$也可能是还不可交易的新工具．(3)市场可能对 $C(t,T)$进行了错误定价．我们可以利用 r_t 的无套利模型计算该工具的公平市场价格，确定合适的对冲、套利或投机头寸．或者，使用投资银行业务中的价格.

转化为PDE(见第21章)，或利用树形图对它进行计算．这都是可能的，因为我们已经得到了 r_t 的动态系统：

$$dr = \tilde{a}(r_t, t)dt + b(r_t, t)dW_t$$

剩下的工作仅仅是计算．

2. **双因子模型**

如果我们要定价的衍生品同时受 r_t 和 R_t(长期利率，与 r_t 不完全相关)影响，那么计算将会变得更复杂．假设该工具的价格记为 $C(r_t, R_t, t)$. 到期日仍为 T，到期收益由已知函数$G(r_T, R_T, T)$确定：

$$C(r_t, R_t, t) = G(r_T, R_T, T)$$

我们仍然可以写出定价公式：

$$C(r_t, R_t, t) = E^Q[e^{-\int_t^T r_s ds} G(r_T, R_T, T)]$$

但是这个模型并不完整．我们还没有得到关于第二个因子 R_t 的无套利模型．在计算该工具的价格之前，我们还需要得到 R_T 风险调整后的SDE. 就这些问题而言，我们向读者推荐Brennan和Schwarz(1979)及相关文献．要注意的是，r_t 与 R_t 之间可能存在复杂的时变相关性，因此计算可能比单因子情形困难得多．[⊖]

3. **调整的重要性**

理解(2.4.1)中即期利率模型的推导过程非常重要．如果我们仅使用计量经济方法估计连续时间漂移系数 $a(r_t, t)$和扩散系数 $\sigma(r_t, t)$，那么得到的模型为

$$dr_t = a(r_t, t)dt + \sigma(r_t, t)dW_t^*$$

我们不能认为该模型是无套利的．计量经济方法得到的是对*真实世界中参数的估计*，该模型在真实测度 P 下是有效的．可以根据连续时间回归残差直接估计维纳过程 W_t^*．

利用

$$B(t, T) = E^Q[e^{-\int_t^T r_s ds}]$$

倒推 r_t 过程得到的才是无套利模型，因为该定价公式使用的是 Q 测度．因此无套利即期利率模型不仅涉及估计或调整的问题，它是建立在选择合适的定价模型的基础上的．

19.2.5 与Black-Scholes环境的比较

我们可以看到，对利率敏感型证券定价的经典方法实质上相当于对即期利率建模．我们也可以看到调整工作并不简单，尤其是当贴现债券价格与期限之间不能完全匹配的时候．

更重要的是，如果我们使用经典方法，套利限制是通过匹配原始收益率曲线间接地包含在所得到的模型中的．我们从一系列贴现债券价格或相应收益率出发，试图找到符合期限结构的 r_t 的模型，使得

⊖ 在后面的章节中，我们将采用不同的符号表示风险调整后的漂移项．在提出一些新概念后，我们可以将风险调整后的漂移项写成 $a(r_t, t) - \lambda_t b(r_t, t)$，其中 λ_t 是Girsanov漂移调整项，这里它等于利率风险的市场价格．

$$B(r_t, T) = E^Q[\mathrm{e}^{-\int_t^T r_s \mathrm{d}s}]$$

对任意 T 都成立.

这与本书前半部分讨论的 Black-Scholes 环境中的基本思想完全不同. 在 Black-Scholes 环境中，通过将标的过程的未知漂移项替换为已知即期利率，套利限制是直接、明确地被包含在模型中的，不需要对股票价格过程的漂移项进行建模，只需用(常数)即期利率 r 对它进行替换. 于是 Black-Scholes 方法将问题简化为对波动率建模. 而假设随机过程 S_t 是几何布朗运动又进一步简化了这一问题，除此以外还假设百分比波动率是常数.

320

从这种意义来看，以即期利率模型为基础的经典方法与目前我们所看到的无套利定价法完全不同.

这就引出了下面的问题：对利率敏感型证券定价时，存在更符合 Black-Scholes 基本思想的方法吗? 答案是肯定的，这种方法就是 Heath-Jarrow-Morton(HJM)模型.

19.3 期限结构的 HJM 方法

我们所考察的套利限制是影响期限不同但其他条件都相同的折价债券价格的共同随机过程作用的结果. 如果所有具有期限结构$\{B(t,T)\}$的流动债券都受相同的不可预测的维纳过程 W_t 的影响，那么债券价格之间一定存在一定的相关性，因为定价关系均为：

$$B(t,T) = E^Q[\mathrm{e}^{-\int_t^T r_s \mathrm{d}s}]$$

对利率敏感型证券定价的经典方法从 $B(t,T)$倒推得到套利关系，并将它们表示为一个无套利即期利率模型：

$$\mathrm{d}r_t = \tilde{a}(r_t, t)\mathrm{d}t + b(r_t, t)\mathrm{d}W_t^*$$

这是一项复杂的任务，间接涉及了市场价格之间各种复杂的套利关系. Heath-Jarrow-Morton(HJM)方法则通过直接引入远期利率对这些套利限制进行了挑战.

这种方法的基础是第 18 章中第二种套利关系. 我们前面提到过，(除期限外完全相同的)折价债券与远期利率之间存在着某种直接的关系. 回顾一个简单情形就足够了.

$B(t,T)$与 $B(t,U)$为两种除期限($T<U$)外完全相同的无违约风险零息债券. $F(t,T,U)$为 t 时刻订立的无违约风险的远期贷款利率，它开始于 T 时刻，结束于 U 时刻. 这里 $F(\cdot)$是 $U-T$ 期间的百分比利率，因此不需要日期调整因子. 那么，第 18 章中讨论过的无套利条件可以写成:[⊖]

$$[1 + F(t,T,U)] = \frac{B(t,T)}{B(t,U)}$$

⊖ 我们用一种稍微不同的语言来重复一下套利条件. $B(t,U)$为在稍晚时刻 U 得到的、确定的 1 美元的现值. 它的倒数是现在的 1 美元在 U 时刻的价值. 用该倒数除以 $1+F(t,T,U)$，将 U 时刻的价值转化为 T 时刻的价值：

$$\frac{1}{[1+F(t,T,U)]B(t,U)}$$

它乘以 $B(t,T)$应该变回最初的 1 美元：

$$B(t,T)\frac{1}{[1+F(t,T,U)]B(t,U)} = 1$$

这是由于 $B(t,T)$是 T 时刻 1 美元的现值.

于是含有 $F(t,T,U)$ 的简单表达式中包含两种期限不同的债券．考虑这些变量共同的动态变化．由于债券是可交易资产，我们可以在相应的 SDE 中将漂移系数变为无风险利率 r_t．

321

到这里所有的推导过程都与 Black-Scholes 推导过程相同．但要注意的是，根据上面的套利关系，两种债券在风险中性环境中的动态变化之比可以用单一的远期利率 $F(t,T,U)$ 的运动来刻画．也就是说，只要确定了在风险中性环境下债券的动态变化，远期利率 $F(t,T,U)$ 的 SDE 也就确定了．不需要调整和/或估计任何其他漂移系数，也没必要调整这些系数中的风险．它们都将自然地被包含在远期利率的动态变化中．

换句话说，如果我们决定对远期利率 $F(t,T,U)$ 而不是即期利率 r_t 进行建模，类似于 Black-Scholes 情形，套利关系将直接包含在远期利率的动态变化中．HJM 方法就是以此为基础发展出来的．当然，在这种结构中，我们需要调整波动率，也需要选择准确的远期利率作为定价基础．

19.3.1 选择哪种远期利率

由于套利关系可以被写成不同的形式，我们在这里有很多不同选择．

原始的 HJM 方法对连续复利瞬时远期利率 $F(t,T)$ 进行建模，即利用第 18 章中得到的关系：

$$B(t,T) = e^{-\int_t^T F(t,s)\mathrm{d}s}$$

其中 $F(t,s)$ 是 s 时刻开始、无穷小区间 $\mathrm{d}s$ 后结束的远期贷款利率．

将该套利关系写成：

$$\frac{B(t,T)}{B(t,U)} = e^{-\int_T^U F(t,s)\mathrm{d}s}$$

我们可以得到连续复合瞬时利率 $F(t,T)$ 的套利限制，这一点将在下一节中进行讨论．

但这仅仅是 HJM 可以采用的其中一种方法．还可以选择使用离散的、非无穷小间隔环境下的远期利率．即我们可以选择以 $F(t,T,U)$ 为基础的模型．令 $U=T+\Delta$，利用该关系得到无套利动态模型：

$$[1+F(t,t+\Delta,t+2\Delta)\Delta] = \frac{B(t,T)}{B(t,T+\Delta)}$$

这里 $\Delta>0$ 固定，考虑 $B(t,\ T)$ 与 $B(t,\ T+\Delta)$ 当 t 变化时的联合动态变化．可以在风险中性测度下，或在第 17 章介绍的远期测度下(取决于要定价的工具)对联合动态变化建模．这样得到的是所谓的 BGM 模型，在 Brace 等(1997)中有详细介绍．本章的剩余部分将采用瞬时远期利率 $F(t,\ T)$ 介绍原始的 HJM 方法．

19.3.2 HJM 方法中的无套利动态变化

根据第 19 章中得到的(期限为 T_i 的)无违约风险纯折价债券价格 $B(t,T_i)$，$T_i<T_{\max}$ 与远期利率 $F(t,T)$ 之间的关系，我们有：

$$B(t,T) = \mathrm{e}^{-\int_t^T F(t,u)\mathrm{d}u} \tag{19.13}$$

该表达式中不含期望运算符，因为 $F(t,T)$ 是 t 时刻的远期利率——开始于未来的 $u>t$ 时刻、持续无穷小间隔 $\mathrm{d}u$ 的远期贷款利率．

在下一节中我们令 $B_t = B(t,T)$，并假设期限为 T 的债券服从下面的随机微分方程：

$$\mathrm{d}B_t = \mu(t,T,B_t)B_t\mathrm{d}t + \sigma(t,T,B_t)B_t\mathrm{d}V_t^T \tag{19.14}$$

其中 V_t^T 是在真实概率测度 Q 下的维纳过程．关于该 SDE 我们要强调 3 点．首先，扩散系数用债券的百分比波动率来表达，但不需要服从几何 SDE 的形式．[㊀] 其次，该 SDE 由指标为 T 的维纳过程驱动．这意味着期限不同的债券通常受不同冲击的影响．我们后面将看到在单因子情况中，所有的 V_t^T 都相同．最后，注意我们书写扩散系数的新形式．$\sigma(t,T,B_t)$ 是关于期限 T 的确切的函数．这一点在下面的推导过程中是必要的，但在后面的章节中可以忽略．

322

债券是可交易资产．在风险中性世界中应用 Girsanov 定理，可以像在 Black-Scholes 环境中一样调整漂移系数：

$$\mathrm{d}B_t = r_tB_t\mathrm{d}t + \sigma(t,T,B_t)B_t\mathrm{d}W_t^T \tag{19.15}$$

其中 r_t 是无风险的瞬时即期利率，W_t^T 是在风险中性测度 Q 下新的维纳过程．即通过从 P 测度变换到 Q 测度，我们消除了债券方程中的未知漂移项．

已知债券的 SDE，我们可以利用(19.13)得到 $F(t,T)$ 的动态变化．从第 19 章得到的套利关系出发：

$$F(t,T,T+\Delta) = \frac{\log B(t,T) - \log B(t,T+\Delta)}{T+\Delta-T} \tag{19.16}$$

其中 $0<\Delta$ 是非无穷小区间，用于定义在 T 时刻开始、在 $T+\Delta$ 时刻结束的非瞬时远期利率 $F(t,T,T+\Delta)$．如果两种债券在其他方面都相同，仅期限不同——相差 Δ，那么它们满足这一假设．

为得到远期利率的无套利动态原理，对(19.16)右边部分应用伊藤引理，并引入风险调整后的漂移项．[㊁] 先对 $\log B(t,T)$ 应用伊藤引理，得到：

$$\mathrm{d}[\log B(t,T)] = \frac{1}{B(t,T)}\mathrm{d}B(t,T) - \frac{1}{2B(t,T)^2}\sigma(t,T,B_t)^2B(t,T)^2\mathrm{d}t \tag{19.17}$$

化简，并带入(19.15)有：

$$\mathrm{d}[\log B(t,T)] = \left(r_t - \frac{1}{2}\sigma(t,T,B_t)^2\right)\mathrm{d}t + \sigma(t,T,B_t)\mathrm{d}W_t \tag{19.18}$$

再对 $\log B(t,T+\Delta)$ 应用伊藤引理，得到的表达式与上面类似，仅用 $T+\Delta$ 替换了上面的 T：[㊂]

㊀ 几何布朗运动 SDE 的扩散系数为 σB_t，其中 σ 是常数．这里我们令 $\sigma(t,T,B_t)$ 也依赖于 B_t．因此这里的百分比波动率不是常数．

㊁ 应用伊藤引理意味着参数 t 发生变化．这里可能有读者会误认为我们在试图求当 $\Delta\to 0$ 时的极限．我们会在后面进行这样的操作，但是在这里 Δ 是常数．

㊂ 毕竟两种债券除期限外完全相同．

$$\mathrm{d}[\log B(t,T+\Delta)] = \left(r_t - \frac{1}{2}\sigma(t,T+\Delta,B_t)^2\right)\mathrm{d}t + \sigma(t,T+\Delta,B_t)\mathrm{d}W_t \quad (19.19)$$

$B(t,T)$和$B(t,T+\Delta)$的SDE中漂移系数的第一项相同，这是因为我们考虑的都是无套利情况，这一点很重要．在Q测度下，期限不同的折价债券预期收益率将等于无风险利率r_t．这一点从本质上来说，与Black-Scholes推导中使用无风险(常)利率r作为股票价格S_t的SDE漂移项相同．

现在将随机微分(19.18)和(19.19)代入$F(t,T,T+\Delta)$的定义式(19.16)中，消去相同的$r_t\mathrm{d}t$项：

323

$$\mathrm{d}F(t,T,T+\Delta) = \frac{1}{2\Delta}[\sigma(t,T+\Delta,B(t,T+\Delta))^2 - \sigma(t,T,B(t,T))^2]\mathrm{d}t + \frac{1}{\Delta}[\sigma(t,T+\Delta,B(t,T+\Delta)) - \sigma(t,T,B(t,T))]\mathrm{d}W_t \quad (19.20)$$

这是对(19.16)应用伊藤引理的最终结果．该等式给出了(开始于T时刻、结束于Δ间隔后的)远期贷款利率的无套利动态原理．

现在我们令$\Delta\to 0$. 得到瞬时远期利率的动态原理．这里注意(19.20)表达式的写法．右边两项的形式类似于：

$$\frac{g(x+\Delta)-g(x)}{\Delta}$$

对于类似这种的表达式，令$\Delta\to 0$意味着取$g(\cdot)$关于x的(标准)导数．将这些项分别写进括号中并令$\Delta\to 0$相当于对右边的两项关于T取导数．得到

$$\lim_{\Delta\to 0}\frac{1}{2\Delta}[\sigma(t,T+\Delta,B(t,T+\Delta))^2 - \sigma(t,T,B(t,T))^2] = \sigma(t,T,B(t,T))\left[\frac{\partial\sigma(t,T,B(t,T))}{\partial T}\right]$$

$$\lim_{\Delta\to 0}\frac{1}{\Delta}[\sigma(t,T+\Delta,B(t,T+\Delta)) - \sigma(t,T,B(t,T))] = \left[\frac{\partial\sigma(t,T,B(t,T))}{\partial T}\right]$$

将它们带入(19.20)，得到瞬时远期利率的SDE：

$$\lim_{\Delta\to 0}\mathrm{d}F(t,T,T+\Delta) = \mathrm{d}F(t,T)$$

或

$$\mathrm{d}F(t,T) = \sigma(t,T,B(t,T))\left[\frac{\partial\sigma(t,T,B(t,T))}{\partial T}\right]\mathrm{d}t + \left[\frac{\partial\sigma(t,T,B(t,T))}{\partial T}\right]\mathrm{d}W_t \quad (19.21)$$

其中$\sigma(\cdot)$为债券价格波动率．

对于该结果我们有下面一些注解．

19.3.3 解释

HJM方法的基础是直接对远期利率进行无套利限制．首先，利用套利关系得到远期利率与债券价格之间的关系．接下来可以写出$B(t,T)$的无套利动态原理．已知债券价格的SDE，可以得到瞬时远期利率应满足的SDE．为看清这其中的真实含义，假设我们给出瞬

时远期利率 $F(t,T)$SDE 的一般形式：

$$dF(t,T) = a(F(t,T),T)dt + b(F(t,T),T)dW_t \tag{19.22}$$

其中 $a(F(t,T),T)$和 $b(F(t,T),T)$是风险调整后的漂移系数和扩散系数，W_t 是风险中性概率下的维纳过程．

读者可能会好奇，如何得到这些在无套利条件下有效的风险调整后的系数呢．前一节的内容说明，在无套利条件下风险调整后的漂移系数可以替换为：

$$a(F(t,T),t) = \sigma(t,T,B(t,T))\left[\frac{\partial\sigma(t,T,B(t,T))}{\partial T}\right] \tag{19.23}$$

扩散项系数可以为：

$$b(F(t,T),t) = \left[\frac{\partial\sigma(t,T,B(t,T))}{\partial T}\right] \tag{19.24}$$

324

因此，前一节得出了瞬时远期利率动态原理中漂移系数的精确无套利限制．这一点与本书前半部分多次使用的 Black-Scholes 方法很相似．在 Black-Scholes 方法中，股票价格 S_t 的 SDE 中，漂移项 μ 被替换为无套利概率测度下的无风险利率 r. 这里的漂移项不替换为 r，而是替换为更复杂的、依赖于债券波动率的项．但该漂移系数通常由套利讨论所决定，仅在远期市场和债券价格之间不存在套利可能时成立．在该过程中并不需对远期利率进行具体建模．

需要强调的是，风险调整后的瞬时远期利率仅依赖于波动率系数．这一点也与 Black-Scholes 环境(不需对标的股票的预期收益率建模)相似，但都需要对波动率建模或调整．在这个意义下，可以认为 HJM 方法是 Black-Scholes 方法在固定收益领域的延伸．

19.3.4 HJM 方法中的 r_t

进一步地，我们注意到 HJM 方法不需对任何短期利率过程建模，也不需要即期利率 r_t 的精确模型．然而，假设市场上存在即期利率．得到的远期利率 $F(t,T)$的 SDE 对即期利率有什么含义呢？这个问题显然是有意义的，因为即期利率等于开始于 t 时刻、期限无穷小的远期贷款利率．

因此对任意 t 都有

$$r_t = F(t,t) \tag{19.25}$$

从远期利率的 SDE 出发，我们可以得到即期利率的相应方程．在开始讨论前，我们先对符号进行简化，将(19.24)中 $b(F(s,T),t)$写成 $b(s,t)$. 然后利用 $b(\cdot)$的新符号，将(19.23)、(19.24)代入(19.21)，写出 $F(t,T)$的积分等式：

$$F(t,T) = F(0,T) + \int_0^t b(s,T)\left[\int_s^T b(s,u)du\right]ds + \int_0^t b(s,T)dW_s$$

接下来，令 $T=t$，得到即期利率 r_t 的值：

$$r_t = F(0,t) + \int_0^t b(s,t)\left[\int_s^t b(s,u)du\right]ds + \int_0^t b(s,t)dW_s \tag{19.26}$$

其中 $b(s,t)$为 $F(s,t)$的波动率．

我们从该等式中得到的第一个重要结果是，在无风险测度下，远期利率是未来即期利率的有偏估计．考虑起始时刻 $t<\tau$ 的未来即期利率 r_τ 的条件期望：

$$E^Q[r_\tau]=E^Q[F(t,\tau)]+E^Q\left[\int_t^\tau b(s,\tau)\left[\int_s^\tau b(s,u)\mathrm{d}u\right]\mathrm{d}s\right]+E^Q\left[\int_t^\tau b(s,\tau)\mathrm{d}W_s\right] \tag{19.27}$$

其中第一个期望值中的远期利率在 t 时刻已知，因此可以去掉它的期望运算符．右边第三个期望值等于 0，因为它是维纳过程的期望．但第二个期望值通常为正，且不为 0. 因此我们有：

$$F(t,\tau)\neq E^Q[r_\tau] \tag{19.28}$$

r_t 的 SDE 的第二个主要含义与即期利率的非马尔可夫性质有关．为了说明这一点，注意(19.26)给出的 r_t 与下面的项相关：

325

$$\int_0^t b(s,t)\left[\int_s^t b(s,u)\mathrm{d}u\right]\mathrm{d}s \tag{19.29}$$

通常来说它是所有过去远期利率波动率的复杂函数．特别地，这一项不仅是把漂移或扩散项过去变化做简单的“累积”：

$$\int_0^t \mu(r_s,\tau)\mathrm{d}s \tag{19.30}$$

或

$$\int_0^t b(r_s,t)\mathrm{d}W_s \tag{19.31}$$

r_t 等式中新出现的项更像是一种交叉乘积．因此 t 时刻前 Δ 时观察到的利率具有相似结构

$$\left[\int_0^{t-\Delta} b(s,t-\Delta)\int_s^{t-\Delta} b(s,u)\mathrm{d}u\right]\mathrm{d}s\right] \tag{19.32}$$

它无法由一种状态变量所捕捉．(19.29)与(19.32)之差为 $t-\Delta$ 前观察到的利率．这说明利率通常不是马尔可夫过程．

下面我们来看一个例子．

常数远期波动率

假设所有远期利率 $F(t,T)$ 的波动率都为常数 b. 则在无套利测度下，任意远期利率都满足方程：

$$\mathrm{d}F(t,T)=b^2(T-t)\mathrm{d}t+b\mathrm{d}W_t \tag{19.33}$$

债券价格的动态变化表示为

$$\mathrm{d}B(t,T)=r_tB(t,T)\mathrm{d}t+b(T-t)B(t,T)\mathrm{d}W_t \tag{19.34}$$

根据这两个等式，我们可以计算(19.26)中的积分，求出即期利率：

$$r_t=F(0,t)+\frac{1}{2}b^2t^2+bW_t \tag{19.35}$$

对应 SDE 为

$$\mathrm{d}r_t=(F_t(0,t)+b^2t)\mathrm{d}t+b\mathrm{d}W_t \tag{19.36}$$

其中 $F_t(0,t)$ 为

$$F_t(0,t) = \frac{\partial F(0,t)}{\partial t} \tag{19.37}$$

注意，根据该模型，即期利率的漂移项与时间相关，波动率为常数．

19.3.5 HJM 方法的其他优点

HJM 方法利用远期利率与债券价格之间的套利关系，直接对瞬时远期利率施加了限制．这样做不需要对即期利率的预期变化率进行建模．

这种方法还有其他优点．如我们在前面几章中看到的那样，k 维马尔可夫过程得到的通常是非马尔可夫单变量模型．因此在 HJM 框架中，我们通常可以令一系列远期利率具有马尔可夫性质，在多变量意义下这是一种合理的近似．然而，当我们在单变量意义下对即期利率进行建模时，后者的变化不具有马尔可夫性质．

这一点很重要，因为当前的实证研究说明现实中的即期利率不具有马尔可夫性质．因此，从这个角度来看，HJM 方法为市场从业者提供了重要的灵活度．

19.3.6 市场实践

从无套利定价的角度来看，显然 HJM 方法更合适．它直接在模型中包含了套利限制，且更灵活．

然而，似乎市场实践仍更倾向于使用经典方法，仍使用即期利率模型．我们如何解释这种矛盾呢?

如 Musiela 和 Rutkowski(1997)中讨论的那样，对瞬时即期利率建模有一定的困难．如果我们将高斯过程代入确定 $dF(t,T)$ 动态原理的 SDE，且使用常数百分比波动率时，该过程就会在时间趋于无穷大时也趋于无穷大．显然这并不是我们希望的动态模型的性质．这会在定价过程中引入不确定性． 326

另外，不论是时间方面还是金钱方面，我们都已经在即期利率模型上投入了大量资源．人们更熟悉即期利率模型，毕竟它对无套利价格有良好的近似效果．

有一些最新的模型利用了*远期测度*来进行建模，这似乎是瞬时远期利率建模问题的一种解决方案，因此也可以考虑用它进行替代．

19.4 如何使 r_t 与初始期限结构相适应

我们在本章中提到过如何使 r_t 适应于已知无套利的期限结构．但我们从未在这些讨论中说明如何在实践中进行操作．本书力图将数值问题的讨论降到最低限度，但也存在一些特殊情况，讨论实际定价模型将有助于理解概念问题，一些计算无套利即期利率模型的简单例子就可以起到这一作用．我们在本章的末尾对这一点进行简要讨论．

假设已知 n 支债券的无套利价格 $B(t,T_i)$，$i=1,\cdots,n$. 同时假设我们使用经典方法对利率敏感型证券进行定价．对于单因子模型，我们首先将风险调整后即期利率模型

$$dr_t = a(r_t, t)dt + b(r_t, t)dW_t$$

适应于该期限结构．如何在实践中进行这一操作呢？

有几种方法可供选择．所有的方法都是从某类合理的即期利率模型出发，然后将它们离散化．因此我们可以令 r_t 满足 Vasicek 模型：

$$dr_t = \alpha(\kappa - r_t)dt + \sigma dW_t$$

接下来直接使用欧拉方法将它离散化：[⊖]

$$r_t = r_{t-\Delta} + \alpha(\kappa - r_t)\Delta + \sigma(W_t - W_{t-\Delta}) \tag{19.38}$$

其中 Δ 是离散区间．剩余的工作取决于所使用的方法．我们来对一些例子进行讨论．

19.4.1 蒙特卡洛方法

假设增量$[W_t - W_{t-\Delta}]$相互独立，且服从均值为 0、方差为 Δ 的正态分布．假设我们已经调整了波动率系数 σ 和均值回归速度 α，这样就只剩下一个未知参数 κ．最后，我们还有初始即期利率 r_0．

考虑下面的问题．利用随机数生成器选择 M 个正态随机变量．将每个随机数乘以$\sqrt{\Delta}$．从对 κ 的历史估计值出发，循环使用方程(19.38)和 r_0，得到第一条蒙特卡洛轨迹 r_t^1．

重复 N 次，得到 N 条这样的即期利率轨迹

$$[\{r_t^1\}, \{r_t^2\}, \cdots, \{r_t^N\}]$$

接下来利用债券定价公式的等价版本计算价格：

$$\hat{B}(t, T_i) = \frac{1}{N}\sum_{j=1}^{N} e^{-\sum_{i=1}^{M} r_t^j \Delta}$$

其中对不同的债券而言 M 可能不同，这取决于相应期限的大小．由于 κ 是随机选择的，因此$\hat{B}(t, T_i)$并不是无套利价格．

但是，我们还有期限结构，它满足无套利条件．因此我们可以尝试调整 κ，使

327

$$\sum_{i=1}^{T_{\max}} |\hat{B}(t, T_i) - B(t, T_i)|^2$$

最小化．这样得到 κ 的值，可以使计算所得的期限结构尽可能地接近观察到的期限结构，在这个意义下，我们可以说 r_t 的动态变化是(近似)无套利的了．

19.4.2 树形模型

我在前面的方法中使用单系数 κ，来确保计算得到的债券价格尽可能接近观察到的期限结构．这种适应并不完美，因为尽管两种期限结构之间的距离会减小，它不会减小到 0．采用一般的树形方法可以改进这一点．

⊖ 欧拉法利用一阶差分替代微分，属于一阶近似，可能导致较大的累积误差．

考虑 r_t 运动对应的二叉树，我们可以选择合适的相关系数，使得该树形轨迹适应于无套利期限结构和相关波动率．例如，假设我们有 N 种无套利债券的价格，并且已知每种债券$B(t,T_i)$的波动率 σ_i．将 r_t 在第 i 阶段的上下波动分别记为 u_i 和 d_i，使得

$$u_i d_i = 1$$

给定这些限制，该树将会是节点重合的，且我们在任一阶段都有 i 个未知参数．接下来的工作就是根据等式：

$$B(0,T_k) = \frac{1}{N}\sum_{j=1}^{N_k} e^{-\sum_{i=1}^{T_k} r_i^j \Delta}$$

确定 u_i 和 d_i，其中 r_i^j 是第 j 条轨迹上第 i 个元素，N_k 为 T_k 后到期的债券树形轨迹数目．这些轨迹依赖于 u_i，d_i，因此可以用这些等式确定 u_i 和 d_i．为此，我们需要足够多的限制，使得树形中未知参数的总数目等于等式的数目．于是可以通过这些等式确定树形参数．树形模型能准确地适应初始期限结构．Black 等(1990)给出了这种方法的一个例子．

19.4.3 封闭形式的解

假设我们可以用解析方法计算数学期望

$$B(t,T) = E^Q[e^{-\int_t^T r_s ds}]$$

并得到 $B(t,T)$封闭形式的解(我们在下一章中进行讨论)．假设得到的结果为函数：

$$B(t,T) = G(r_t, T, \kappa)$$

下面我们可以通过选择 κ，最小化该解与观察到的无套利收益率曲线之间的距离：

$$\sum_{i=1}^{T_{\max}} |B(t,T_i) - G(r_t, T, \kappa)|^2$$

这是得到 r_t(近似)无套利模型的另外一个例子．

19.5 小结

本章简要总结了利率敏感型衍生证券的两种主要定价方法．经典方法对即期利率进行建模，在适应原始曲线的过程间接包含了套利限制．而 HJM 方法则是对 Black-Scholes 公式在利率敏感型证券定价方面的推广． 328

19.6 参考阅读

该问题最优秀的参考书目是 Musiela 和 Rutkowski(1998)．当然，它技术性非常强，但我们还是希望那些对固定收益证券领域感兴趣的读者尽量阅读并熟悉它．Jarrow(1996)对离散时间情况精彩的处理也值得推荐．

19.7 习题

1. 考虑下面等式，它给出了将时间$[0,T]$分成 n 等份(每份长度均为 Δ)时，利率的动态原理：

$$r_{t+\Delta} = r_t + \alpha r_t + \sigma_1(W_{t+\Delta} - W_t) + \sigma_2(W_t - W_{t-\Delta})$$

其中随机误差项

$$\Delta W_t = (W_{t+\Delta} - W_t)$$

服从正态分布

$$\Delta W_t \sim \mathcal{N}(0,\sqrt{\Delta})$$

(a) 解释该等式中误差项的结构．特别地，你认为 $\Delta W_{t-\Delta}$ 进入利率动态原理合理吗？

(b) 可以写出该公式在连续时间条件下的随机微分方程吗？会遇到什么困难？

(c) 假设已知长期利率 R_t 的动态原理为

$$R_{t+\Delta} = R_t + \beta r_t + \theta_1(\widetilde{W}_{t+\Delta} - \widetilde{W}_t) + \theta_2(\widetilde{W}_t - \widetilde{W}_{t-\Delta})$$

其中协方差

$$E[\Delta\widetilde{W}\Delta\widetilde{W}] = \rho\Delta$$

请写出向量过程

$$X_t = \begin{bmatrix} r_t \\ R_t \end{bmatrix}$$

的表达式，使得 X_t 是一阶马尔可夫过程．

(d) 可以写出该系统在连续时间条件下的等价形式吗？

(e) 假设短期利率或长期利率不单独具有马尔可夫性．它们的向量有可能具有马尔可夫性吗？

2. 假设(向量)马尔可夫过程 X_t

$$X_t = \begin{bmatrix} r_t \\ R_t \end{bmatrix}$$

具有下面的动态原理

$$\begin{bmatrix} r_{t+\Delta} \\ R_{t+\Delta} \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \begin{bmatrix} r_t \\ R_t \end{bmatrix} + \begin{bmatrix} \Delta W^1_{t+\Delta} \\ \Delta W^2_{t+\Delta} \end{bmatrix}$$

其中误差项向量服从正态分布但无序列相关性．假设 r_t 是短期利率，R_t 是长期利率．

(a) 导出短期利率 r_t 的单变量表达式．

(b) 根据该表达式，r_t 是马尔可夫过程吗？

(c) 单变量过程 r_t 在什么条件下是马尔可夫过程？

3. 假设当 $t=0$ 时，已知 4 种分别在 $t=1,2,3,4$ 时刻到期的债券价格为 $\{B_1,B_2,B_3,B_4\}$．这是利率的期限结构．已知一期的远期利率 $\{f_0,f_1,f_{02},f_3\}$，其中 f_i 为 $t=0$ 时刻订立、开始于 $t=i$ 时刻、结束于 $t=i+1$ 时刻的远期贷款．换句话说，如果借方在 $t=i$ 时借出 N 美元，他或她需要在 $t=i+1$ 时归还 $N(1+f_i)$ 美元．用 r_i 表示即期利率．根据定义我们有

329

$$r_0 = f_0$$

$\{B_i\}$ 及所有远期贷款均无违约风险．在任一时点都只可能有两种可能的状态，记为 $\{u_i,d_i;i=1,2,3,4\}$．

(a) 从 $i=0$ 时来看，$i=3$ 时可能存在多少种状态？

(b) 假设 $\{B_1=0.9,\ B_2=0.87,\ B_3=0.82,\ B_4=0.75\}$，且 $\{f_0=8\%,\ f_1=9\%,\ f_2=10\%,\ f_3=18\%\}$．构造 3 个套利组合，使得 $i=1,\ 2,\ 3$ 时无风险收益大于零．

(c) 构造 3 个套利组合，使得 $i=0$ 时刻无风险收益大于零．

(d) 已知无风险零息债券 B_n 在 $t=n$ 时刻到期，远期利率为 $\{f_0,\cdots,f_{n-1}\}$．求 B_n 关于 f_i 的函数．

(e) 考虑将金融基本定理应用到该系统中：

$$\begin{bmatrix} B_1 \\ B_2 \\ B_3 \\ B_4 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ B_2^u & B_2^d \\ B_3^u & B_3^d \\ B_4^u & B_4^d \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix}$$

所有的 B_i 可以独立确定吗？

(f) 在上面的系统中，所有的$\{f_i\}$可以独立确定吗？

(g) 可以认为所有的 f_i 都服从正态分布吗？请给出证明．

4. 再次考虑第 1 题的环境．假设我们要对 3 种欧式看涨期权进行定价，期权标的为一期（即期）Libor 利率 L_i，$i=0$，1，2，3．将这些期权价格记为 C_i．每种期权的收益均为：

$$C^i = N\max[L_i - K, 0]$$

其中 N 为名义本金，不失一般性，我们可以令它等于 1．

(a) 如何对这种期权进行定价？

(b) 有如下假设：

i. 每个 f_i 是对于未来未知量 L_i 的现在的观察值．

ii. 每个 f_i 都服从正态分布，且均值为 0，方差为 σ_i．

iii. 可以使用 Black 公式对看涨期权定价．

(c) 在风险中性测度下，利用这些假设对货币市场账户正规化合适吗？请解释．

(d) 对应于每个 L_i 的远期测度可以对该问题有所改善吗？

(e) 可以得到 $t=1$，2 时的远期测度吗？

(f) 利用远期测度，对 $t=2$ 时的看涨期权定价．

5. 考虑开始于固定时间 T_n 的远期互换，它在 $T_M > T_n$ 时到期．假设应计期限长度按年度计为 δ．由于支付是延迟的，第一笔支付发生在 $T_{n+1} = T_n + \delta$，最后一笔支付发生在 T_{M+1}．首先，$t < T_n$ 时该远期互换 SW_t 的值为

$$SW_t = E\left(\delta \sum_{j=n}^{M} \frac{B_t}{B_{T_{j+1}}}(L(T_j, T_{j+1}) - R)\right) \tag{19.39}$$

其中 R 为合约中确定的固定（年化）利率，$L(T_j, T_{j+1})$是区间$[T_j, T_{j+1}] = [T_j, T_j + \delta]$对应的即期 Libor 利率．

注意 Libor 利率是年化单利利率，也就是说我们在 T_j 时刻投资的 1 美元，到 $T_j + \delta$ 时按 Libor 利率 $L(T_j, T_{j+1})$计算价值将为 $1 + \delta L(T_j, T_{j+1})$．

假设名义本金为 1 美元，证明

$$SW_{T_n} = 1 - Z_{T_n}^{T_{M+1}} - R\delta \sum_{j=n+1}^{M+1} Z_{T_n}^{T_j} \tag{19.40}$$

6. 考虑下面即期利率的随机微分方程：

$$dr_t = \lambda(\theta - r_t)dt + \sigma r_t dW_t \tag{19.41}$$

使用下面的参数：

$$r_0 = 0.03 \tag{19.42}$$

$$\lambda = 1.5 \tag{19.43}$$

$$\theta = 0.06 \tag{19.44}$$

$$\sigma = 0.25 \tag{19.45}$$

$$T = 1 \tag{19.46}$$

编写程序，通过拟合得出一列短期利率．利用这些短期利率计算债券价格的期望值和方差．

330
~
331

第 20 章　利率衍生品的经典 PDE 分析

20.1　引言

读者们对于 Black-Scholes 公式的各种推导方法(包括偏微分方程法)已经很熟悉了．特别地，第 12 章中证明了如何将无风险借贷、标的工具及相应期权结合起来，构造无风险投资组合．这种投资组合持有头寸发生的微小随机扰动可以相互抵消，使得该投资组合的收益是*确定的*．于是在无违约风险的情况下，该投资组合的收益率等于无风险即期利率 r（我们认为 r 是常数)．否则将存在套利机会．在这个条件下应用伊藤公式，得到 Black-Scholes 的基础 PDE. Black-Scholes 的 PDE 形式为：

$$-rF+F_t+rS_tF_s+\frac{1}{2}\sigma^2S_t^2F_{ss}=0 \tag{20.1}$$

其边界条件为：

333

$$F(S_T,T)=\max[S_T-K,0] \tag{20.2}$$

其中 r 无风险瞬时即期利率是常数，S_t 是(不支付分红的)股票价格，F 是以该股票为标的的欧式看涨期权在 t 时刻的价格．K 和 T 分别是该看涨期权的执行价格和到期日．我们在第 15 章中提到过，该 PDE 的解等于在下面风险中性的 Q 测度下的条件期望：

$$F(S_t,t)=E_t^Q[\mathrm{e}^{-r(T-t)}F(S_T,T)] \tag{20.3}$$

假设我们现在讨论的是以利率敏感型证券为标的的衍生品，我们(至少)会问两个问题：

- 对于利率衍生品，我们可以得到类似的 PDE 吗？比如考虑最简单的情况，无违约风险贴现债券的价格满足哪种 PDE?
- 给定利率衍生品对应的 PDE，我们可以得到其类似于(20.3)的形式为条件期望的解吗?

这些问题可以通过两种方法进行回答．第一种方法是，我们可以采用第 12 章中的方法，按照与经典 Black-Scholes PDE 类似的推导方式，来得到折价债券价格的 PDE. 特别地，我们可以构造一个无风险投资组合，令它的确定收益等于对储蓄账户的无风险瞬时投资．应用伊藤公式可以得到合意的 PDE. ⊖

第二种得到利率敏感型证券 PDE 的方法是利用鞅等价并直接应用所谓的 Feynman-Kac 公式．事实上，当我们考察特定类别的期望与 PDE 之间的关系时，我们会得到一个有趣的数学规律．可以证明：

$$B(t,t)=E_t^Q[\mathrm{e}^{-r\int_t^T r_s\mathrm{d}s}B(T,T)] \tag{20.4}$$

⊖ 这里我们要提醒读者，无风险投资组合不是自融资组合，因此从数学角度看，该方法在连续时间环境下不是很精确．但我们仍可以得到“正确的”PDE，因为额外投资或收回的现金流期望值为 0. 该问题在第 12 章中有更详细的讨论．我们继续使用这种方法，但读者要注意相应的条件．

等表达式与特定类别的偏微分方程之间具有紧密的联系. 在随机微积分中，这些问题属于"伊藤扩散过程的无穷小生成算子""Kolmogorov 后向方程"以及更为重要的"Feynman-Kac公式"的内容. 给定类似于(20.4)的条件期望，我们可以用这些方法直接得到它对应的PDE，反之亦然. 在理论上，我们还需要一些关于所涉及的随机变量的其他条件，才能保证这种对应关系成立，但对于金融市场从业者来说，这仍然是便利的工具. 我们将会在下一章对这些现代工具(随机分析)进行讨论.

本章使用经典的方法来说明利率衍生品的价格满足与基础 Black-Scholes PDE 相似的PDE. 但是该推导过程将与第 12 章完全不同，因为现在的基础变量是即期利率 r_t，而即期利率不同于 S_t(Black-Scholes 环境中可交易资产价格)，它不是资产价格.[㊀] 显然，这里也会出现即期利率建模方面的困难.

利率衍敏感型证券基础 PDE 的推导过程与 Vasicek(1977)经典论文中的步骤类似. 基本思想是将决定多种资产随机运动的单一维纳过程中隐含的套利条件嵌入到收益率动态过程中.[㊁] 在 Black-Scholes 方法中，我们涉及的是标的股票及其对应的看涨期权两种证券. 股票价格的无穷小随机波动也会影响期权价格. 因此同一种随机波动会影响两种价格. 这两种证券可以与无风险借贷小心地结合起来，使不可预测的随机波动相互抵消，得到一个无风险的投资组合.

334

这种思想也适用于利率敏感型证券. 例如，除期限外，不同的债券其实都是相似的：它们都受相似的无穷小随机波动影响. 因此，在某些条件下，只要仔细选择投资组合的权重，一个包含两种(或更多)债券的投资组合可以是无风险的.

但是与股票相比，还是有一些差别的. 在经典 Black-Scholes 推导中，我们认为即期利率是常数. 该假设不是很严格. 对于利率敏感型证券来说，常数利率的假设不成立. 相反，驱动整个系统的随机波动就是来自影响瞬时即期利率 r_t 的无穷小维纳增量. 但后者不是前面提到过的资产价格. 不能简单地根据套利讨论，令利率动态变化的未知漂移系数等于无风险利率. 这就导致了利率衍生品 PDE 推导过程和数值计算的复杂性. 事实上，尽管下面的数学推导过程很直观，它们还是要比标的为股票的普通看涨期权复杂的.

最后，我们要重申，这里采用的经典方法只是类似于 Black-Scholes PDE. 从技术角度来看，正确的推导要包含对自融资无风险投资组合的讨论. 如前所述，下面的方法可能无法得到自融资投资组合.

20.2 基本框架

第一步是设定基本框架. 假设描述两种无违约风险折价债券价格 $B(t,T_1)$、$B(t,T_2)$ 动态过程的 SDE 是给定的，期限分别为 T_1 和 T_2 $(T_1<T_2)$. 债券价格由相同的维纳过程 W_t 驱动. 为简化记法，我们在本节中忽略时间下标 t，将债权价格记为：

㊀ r_t 多为百分比收益率，而不是纯粹的数字.

㊁ 在两因子模型中就是两个相互独立的维纳过程.

$$B^1 = B(t, T_1) \tag{20.5}$$

$$B^2 = B(t, T_2) \tag{20.6}$$

假设债券价格的动态变化服从以下规律：

$$dB^1 = \mu(B^1, t)B^1 dt + \sigma_1(B^1, t)B^1 dW_t \tag{20.7}$$

$$dB^2 = \mu(B^2, t)B^2 dt + \sigma_2(B^2, t)B^2 dW_t \tag{20.8}$$

有两点需要注意．第一，扩散项都是 W_t 的函数，但依赖于不同的扩散系数 σ_i，$i=1$，2．第二，波动率系数为百分比波动率，但债券价格变化规律不一定由几何布朗运动给出，因为漂移系数和扩散系数也取决于 B_i，$i=1$，2，而不是几何布朗运动 SDE 要求的常数．

由于我们采用的是经典方法，所以需要一个利率模型．假设 r_t 的动态原理为：

$$dr_t = a(r_t, t)dt + b(r_t, t)dW_t \tag{20.9}$$

335 其中漂移系数 $a(r_t, t)$和扩散系数 $b(r_t, t)$已知．它们由历史数据或实践方法(用市场价格进行调整)得出．要强调的是，这里的 W_t 是在真实世界 P 测度下的维纳过程．

注意该即期利率动态原理的关键限制，我们假设系数 $a(r_t, t)$、$b(r_t, t)$仅依赖于 r_t 最新的观测值，因此之前的 $r_s(s<t)$不影响漂移系数和波动率系数．根据前一章可知，r_t 的这种马尔可夫性质在一般的期限结构模型中会被破坏．经典方法仍假设它是一种合理的近似．

20.3 利率风险的市场价格

要导出贴现债券价格的 PDE，我们首先需要得到 t 时刻下包含两种债券的无风险投资组合$\mathcal{P}$．[⊖]不失一般性，假设买入 θ_1 单位 B^1，卖出 θ_2 单位 B^2，投资组合总价值为：

$$\mathcal{P} = \theta_1 B^1 - \theta_2 B^2 \tag{20.10}$$

假设投资权重为：

$$\theta_1 = \frac{\sigma_2}{B^1(\sigma_2 - \sigma_1)}\mathcal{P} \tag{20.11}$$

$$\theta_2 = \frac{\sigma_1}{B^2(\sigma_2 - \sigma_1)}\mathcal{P} \tag{20.12}$$

其中 $\sigma_i(i=1$，2)为方程(20.7)、(20.8)中两种债券的波动率系数 $\sigma(B^1, t)$、$\sigma(B^2, t)$．该投资组合的价值随时间变化而变化．如果投资组合权重始终不变，那么投资组合价值的无穷小变化将为：

$$d\,\mathcal{P} = \theta_1 dB^1 - \theta_2 dB^2 \tag{20.13}$$

代入 dB^1，dB^2 对应 SDE，有：

$$\begin{aligned} d\,\mathcal{P} =& \theta_1[\mu(B^1, t)B^1 dt + \sigma_1(B^1, t)B^1 dW_t] \\ & - \theta_2[\mu(B^2, t)B^2 dt + \sigma_2(B^2, t)B^2 dW_t] \end{aligned} \tag{20.14}$$

⊖ 为书写简便，省略时间下标．

将维纳增量 $\mathrm{d}W_t$ 移项重组，我们可以看到它的系数在代入 θ_1 和 θ_2 的值后变为 0：

$$(\theta_1\sigma_1 B^1 - \theta_2\sigma_2 B^2) = \left(\frac{\sigma_2}{B^1(\sigma_2-\sigma_1)}\sigma_1 B^1 - \frac{\sigma_1}{B^2(\sigma_2-\sigma_1)}\sigma_2 B^2\right)\mathcal{P} = 0 \tag{20.15}$$

于是投资组合的增量变化为：

$$\mathrm{d}\,\mathcal{P} = (\theta_1\mu_1 B^1 - \theta_2\mu_2 B^2)\mathrm{d}t \tag{20.16}$$

这些增量不包含维纳成分，是完全可预测的．

这些步骤依赖于投资组合权重 θ_1 和 θ_2 的值，这些权重使得投资组合$\mathcal{P}$的 SDE 中没有 $\mathrm{d}W_t$，这一点类似于 Black-Scholes PDE 的推导．事实上，代入 θ_i，除以再乘以$\mathcal{P}$后，可以将 $\mathrm{d}\,\mathcal{P}$写成：

$$\mathrm{d}\,\mathcal{P} = \frac{(\sigma_2\mu_1 - \sigma_1\mu_2)}{(\sigma_2-\sigma_1)}\mathcal{P}\mathrm{d}t \tag{20.17}$$

该 SDE 不包含扩散项，$\mathrm{d}\,\mathcal{P}$的动态行为是无风险的．因此，我们现在可以进行标准化讨论，认为该投资组合不存在任何套利机会，其确定的收益率等于

$$r_t\,\mathcal{P}\,\mathrm{d}t = \frac{(\sigma_2\mu_1 - \sigma_1\mu_2)}{(\sigma_2-\sigma_1)}\mathcal{P}\mathrm{d}t \tag{20.18}$$

336

化简$\mathcal{P}\mathrm{d}t$，移项，得到

$$\frac{\mu_1 - r_t}{\sigma_1} = \frac{\mu_2 - r_t}{\sigma_2} \tag{20.19}$$

也就是说，用相应波动率参数进行标准化后，期限不同的债券提供的风险溢价是相同的．波动率更高的债券按比例支付更高的风险溢价．[⊖]这个结果并不出乎意料，因为这些债券的风险来源相同，都是 $\mathrm{d}W_t$ 因子．显然，如果某一支债券是另外一种维纳过程 W_t^* 的函数，那么即使无套利条件成立，不同债券每单位波动率的风险溢价也不同．注意，这些风险溢价可以为负数．

在该推导过程中，基础债券的期限是任意的，因此只要它们的动态变化受同一维纳过程 W_t 驱动，所有折价债券对应的方程就都是相似的．这就得到了与所有债券价格 $B(t,T_i)$ 都相关的项 $\lambda(r_t,t)$：

$$\frac{\mu_i - r_t}{\sigma_i} = \lambda(r_t,t) \tag{20.20}$$

该项被称为利率风险的市场价格．从推导过程中可以看出，它通常是 r_t 和 t 的函数．在下一节中，我们将它简写为 λ_t．再次注意，λ_t 与债券期限相互独立．

值得一提的是，Black-Scholes 框架中也出现了类似的权益风险市场价格，但没有直接使用它．相比于 Black-Scholes PDE，对利率敏感型证券 PDE 的推导而言，我们需要使用 λ_t．

⊖ 另一种说法是这些债券的夏普比率相等．

20.4　PDE 的推导

债券价格 PDE 推导过程的第三步是使用 $B(r_t, T)$伊藤的扩展形式．记住，$B(r_t, T)$也是 r_t 的函数，应用伊藤定理得到：

$$\mathrm{d}B(r_t, t) = B_r \mathrm{d}r_t + B_t \mathrm{d}t + \frac{1}{2} B_{rr} b\,(r_t, t)^2 \mathrm{d}t \tag{20.21}$$

代入 $\mathrm{d}r_t$

$$\mathrm{d}r_t = a(r_t, t)\mathrm{d}t + b(r_t, t)\mathrm{d}W_t \tag{20.22}$$

得到

$$\mathrm{d}B(r_t, t) = \left(B_r a\,(r_t, t) a(r_t, t) + B_t + \frac{1}{2} B_{rr} b\,(r_t, t)^2\right)\mathrm{d}t + b(r_t, t)\mathrm{d}W_t \tag{20.23}$$

其中 W_t 仍为在真实世界概率测度 P 下的维纳过程．该 SDE 与驱动债券价格动态变化的原始方程相同．将记法简化，该 SDE 在 P 测度下为：

$$\mathrm{d}B = \mu(B, t) B\mathrm{d}t + \sigma(B, t) B\mathrm{d}W_t \tag{20.24}$$

这意味着我们可以令两个方程的漂移系数和扩散系数分别相等．令(20.23)和(20.24)中的扩散系数相等，则

$$b(r_t, t) B_r = \sigma B \tag{20.25}$$

其中 $\sigma(B, t)$缩写为 σ. 令(20.23)和(20.24)中的漂移系数相等，则：

$$\mu(B, t) B = B_r a\,(r_t, t) + B_t + \frac{1}{2} B_{rr} b\,(r_t, t)^2 \tag{20.26}$$

于是我们得到可以用于计算债券价格 PDE 的(20.25)和(20.26). 事实上，方程(20.26)已经是 PDE，只是它含有未知项 $\mu(B, t)$. 另外要注意的是，到此为止我们都没有将套利限制包含进来，而它是该系统必须满足的．[⊖]

337

可以证明，使用套利条件就可以将(20.26)中的未知漂移项 $\mu(B, t)$消去．回顾 Black-Scholes PDE 的情形，我们简单地用常即期利率 r 替代 $\mu(B, t)$. 但在目前的情况中，这种处理方法是不可能的，因为我们在(20.26)中使用了即期利率漂移系数 $a(r_t, t)$. 如果我们将$\mu(B, t)$换为 r_t，需要将(20.26)中即期利率的漂移系数 $a(r_t, t)$也调整为在风险中性测度下的等价系数．但 r_t 不是资产价格，我们不清楚如何进行这种调整．而这时，使用利率风险的市场价格 λ_t 可以解决该问题．

事实上，(20.20)给出了利率风险的市场价格 λ_t：

$$\frac{\mu(B, t) - r_t}{\sigma} = \lambda_t \tag{20.27}$$

用(20.25)中扩散系数表示为：

$$\frac{B(\mu(B, t) - r_t)}{b(r_t, t) B_r} = \lambda_t \tag{20.28}$$

⊖ 由于我们假设所有债权价格都受同一个维纳过程 W_t 驱动，因此将存在讨论限制．

于是有：

$$\mu(B,t)B = r_t B + b(r_t,t)\lambda_t B_r \tag{20.29}$$

将 $\mu(B,t)B$ 的右边部分代入(20.26)，移项有：

$$B_r a(r_t,t) + B_t + \frac{1}{2}B_{rr}b(r_t,t)^2 - r_t B - b(r_t,t)B_r\lambda_t = 0 \tag{20.30}$$

注意，未知漂移系数 $\mu(B,t)$ 被消掉了．最终可以写成：

$$B_r(a(r_t,t) - b(r_t,t)\lambda_t) + B_t + \frac{1}{2}B_{rr}b(r_t,t)^2 - r_t B = 0 \tag{20.31}$$

这是无违约风险的纯折价债券价格 $B(t,T)$ 对应的 PDE. 相应边界条件要比 Black-Scholes 环境简单．该债券无违约风险，无论到期时即期利率处于什么水平，债券到期时价值一定为 1：

$$B(T,T) = 1 \tag{20.32}$$

如果我们有利率模型，且其漂移系数 $a(r_t,t)$ 和扩散系数 $b(r_t,t)$ 已知，为了在实践中使用该 PDE，我们仍需要 λ_t 的估计量．否则我们无法使用该方程．此外，值得强调的是该 PDE 中 B_r 的系数等于对风险进行调整后即期利率动态原理的漂移系数．

事实上，我们使用的似乎是在风险中性侧度 Q 下即期利率动态过程中的漂移系数．对(20.9)应用 Girsanov 定理，将定义在 P 测度下的维纳过程 W_t 变为在 Q 测度下定义的维纳过程 $\widetilde{W}_t$，我们可以得到 r_t 新的 SDE：

$$\mathrm{d}r_t = (a(r_t,t) - b(r_t,t)\lambda_t)\mathrm{d}t + b(r_t,t)\mathrm{d}\widetilde{W}t \tag{20.33}$$

而这个 SDE 的漂移系数根据利率风险进行了调整．无论我们是否将债券价格的漂移系数从 $\mu(\cdot)$ 变为 r_t，我们都需要将即期利率动态原理中的 $a(r_t,t)$ 变为 $(a(r_t,t)-b(r_t,t)\lambda_t)$.

现在我们对该推导过程进行总结，将它与 Black-Scholes 环境下的 PDE 进行比较．

比较

推导 PDE 的一般方法类似于 Black-Scholes 的情形．主要差别在于，当前的环境中驱动过程不是资产价格 S_t，而是即期利率 r_t(纯数值). 因此引入无套利条件方式也不同，不能令未知的漂移系数等于无风险利率．

所使用的方式用的是 r_t 的风险市场价格，来对债券价格动态变化中的漂移系数进行修正. 读者应该意识到，(20.29)间接地为该等式引入了无套利条件．

$$\mu(B,t)B = r_t B + (b(r_t,t)B_r)\lambda_t \tag{20.34}$$

338

然而，注意一个非常重要的差别．在进行 Black-Scholes 环境下的推导时，通过无套利条件，完全不需要对股票价格过程 S_t 和漂移系数进行建模和调整．事实上，在 Black-Scholes 推导中，S_t 的期望变化值并不重要．期权价格仅依赖于相关波动率．

用即期利率方法给利率敏感型证券定价时，使用无套利条件仍可以在 PDE 中再次引入即期利率 r_t. 然而在该过程中，出现了两个新参数：即期利率漂移系数 $a(r_t,t)$ 和利率风险的市场价格 λ. 如果要在真实世界中使用该 PDE，需要估计或调整这些参数．我们在前一

章中提到过，这与 Black-Scholes 方法(仅需要对波动率建模)出现了差异，基本思想也有了相应变化，因为我们现在需要 r_t 动态过程的完整模型.

上面推导过程的第二个基本要点在于单一驱动过程 r_t 的假设. 我们假设所有债权价格的动态变化都受相同的单变量维纳过程 W_t 驱动. 相同的维纳过程也出现在即期利率 r_t 的 SDE 中，该假设使我们可以得到便于使用的无套利条件(是利率风险市场价格 λ_t 的函数). 显然，这种情形不会总是成立. 股票价格是单变量随机过程 W_t 的函数这一假设是近似可接受的，但是对于期限从特别短到特别长的一系列折价债券都进行相同的操作就不合适了.

本书的目的主要在于介绍相关工具，而不是真实市场中令人满意的定价方法. 因此我们在这里只在单因子的假设下进行讨论.㊀

20.5 PDE 的封闭形式解

债券价格基础 PDE 有可能有封闭形式的解，由此得到一个将 $B(t,T)$ 与期限 T、当前即期利率 r_t 和相关系数 $a(r_t,t)$、$b(r_t,t)$、λ_t 联系起来的公式.

这里的情形与 Black-Scholes 基础 PDE 和 Black-Scholes 公式类似. 如果对 S_t 过程和常利率进行相关假设，那么解 PDE 就可以得到 Black-Scholes 公式. 在目前的框架中，如果对利率过程 r_t 进行相关假设，我们也可以解出债券价格 PDE. 我们来讨论几个简单的例子.

20.5.1 情形 1：r_t 确定

我们先来讨论一种极端情形. 假设即期利率为常数，对任意 t 有 $r_t=r$. 则 r_t 的 SDE：

$$dr_t = a(r_t,t)dt + b(r_t,t)dW_t$$

将具有如下(平凡)系数：

$$a(r_t,t) = b(r_t,t) = 0$$

另外，由于不存在利率风险，不需支付风险溢价，故：

$$\lambda = 0$$

因此 $B(t,T)$基础 PDE 的原始形式为：

$$B_r(a-b\lambda) + B_t + \frac{1}{2}B_{rr}b^2 - r_tB = 0 \tag{20.35}$$

化简为

$$B_t + rB = 0$$

边界条件为

339

$$B(T,T) = 1$$

㊀ 要记住的是，实际定价过程中也经常用到这个假设.

但这与终端条件为 $B(T)=1$ 的常微分方程相同

$$\frac{\mathrm{d}B(t,T)}{\mathrm{d}t}+rB(t,T)=0$$

该方程的解为

$$B(t,T)=\mathrm{e}^{-r(T-t)}$$

该债券定价公式满足边界条件和基础 PDE. 它就是通常的用常瞬时利率 r 进行贴现的贴现公式.

20.5.2 情形 2：r_t 为均值回归过程

假设风险的市场价格为常数：

$$\lambda(r_t,t)=\lambda \tag{20.36}$$

但即期利率满足均值回归 SDE：

$$\mathrm{d}r_t=\alpha(\kappa-r_t)\mathrm{d}t+b\mathrm{d}W_t \tag{20.37}$$

其中 W_t 是在真实世界测度下的维纳过程. 注意，这里我们假设波动率为常数 b. 进一步假设参数 α，κ，b 和 λ 都是已知的. 则 $B(t,T)$的基础 PDE 将化简为：

$$B_r(\alpha(\kappa-r_t)-b\lambda)+B_t+\frac{1}{2}B_{rr}b^2-r_tB=0 \tag{20.38}$$

这就是著名的 Vasicek 模型. Vasicek(1977)提出了这个模型.

可以证明该 PDE 的闭形式解是债券定价公式，在 $t=0$ 时刻，

$$B(0,T)=\mathrm{e}^{\frac{1}{\alpha}(1-\mathrm{e}^{\alpha T})(R-r)-TR-\frac{b^2}{4\alpha^3}(1-\mathrm{e}^{-\alpha T})^2} \tag{20.39}$$

其中

$$R=\kappa-\frac{b\lambda}{\alpha}-\frac{b^2}{\alpha^2} \tag{20.40}$$

r 是即期利率的当前观测值. 对未知参数进行合理估计后就可以画出该函数的图像.

例子

考虑即期利率长期均值为 5%，且即期利率以 0.25 的速度趋近于长期均值的经济体. 我们有

$$\alpha=0.25 \quad \kappa=0.05 \tag{20.41}$$

除此以外，假设一年内利率波动率为 0.015：

$$b=0.015 \tag{20.42}$$

要应用该公式，我们需要利率风险的市场价格. 假设有

$$\mu-r_t=-0.1\sigma \tag{20.43}$$

其中 μ，σ 是未知的债券漂移参数和波动率参数. 则有

$$\lambda=-0.10 \tag{20.44}$$

利用这些参数，我们可以计算出债券定价函数 $B(t,T)$，它是原始利率 r 和期限 T 的函数. 这就是前面讨论的所谓的贴现曲线.

当$\{\lambda=-0.10，b=0.015，\alpha=0.25，\kappa=0.05\}$时，$\{B(t,T)，T\in[0,T_{\max}]\}$在即期利

率分别为 $r=0.5\%$，$r=5\%$，$r=15\%$时的图像如图 20-1 所示．由于它们是折价债券，短期债券的价值接近 1，期限越长，价格越低．

对贴现曲线取对数再除以期限可以得到相应收益率曲线．图 20-2 给出的是上述情形的收益率曲线．

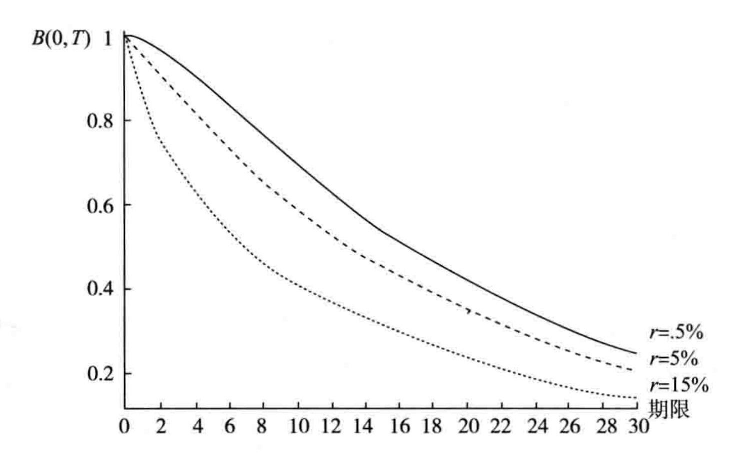


图 20-1 当即期利率分别为 $r=0.5\%$，$r=5\%$，$r=15\%$时零息债券的图像

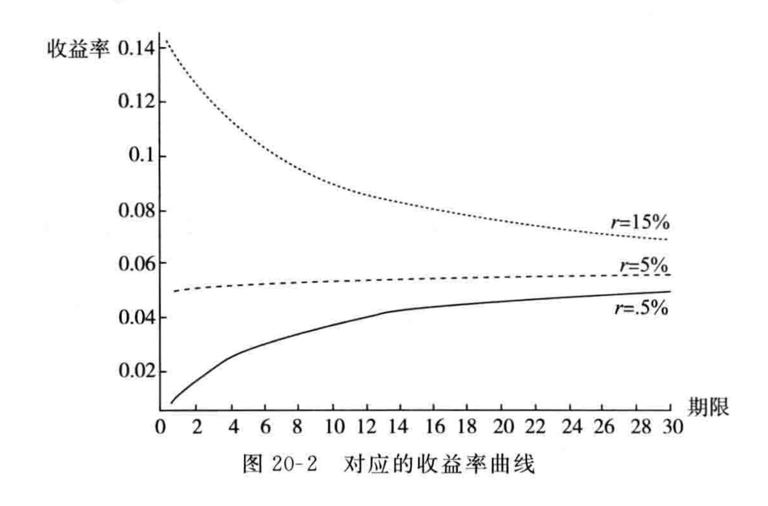


图 20-2 对应的收益率曲线

注意，利率 SDE 的均值回归特征决定了收益曲线是向上倾斜、向下倾斜还是平坦的．340 因为如果当前的即期利率为 15%，那么该模型会假设它向 5%的均值运动．因此长期债券就可以用 5%的平均利率来进行定价，而短期债券要用 15%左右的利率来进行定价．当前的短期低于长期均值后出现的就是相反的情形，得到向上倾斜的收益率曲线．

图 20-3 假设 $r=5\%$，对风险市场价格 λ 变化对贴现曲线的影响进行了说明．

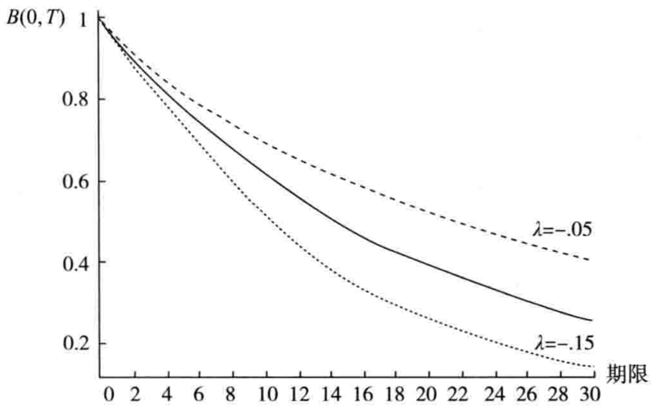


图 20-3 当 $r=5\%$时风险市场价格 λ 对贴现曲线的影响

20.5.3 情形 3：更复杂的形式

还有几个可以导出债券价格封闭形式解的模型．

例如，在 Cox-Ingersoll-Ross 模型中，假设基础即期利率 r_t 满足一个略微有所不同的 SDE [341]

$$dr_t = \alpha(\kappa - r_t)dt + br^{\frac{1}{2}}dW_t \tag{20.45}$$

这就是著名的利率波动率平方根形式的表达．

相应的 PDE 为

$$(\alpha(\kappa - r_t) - b^2 r\lambda) + B_r + B_t + \frac{1}{2}B_{rr}b^2 r - rB = 0 \tag{20.46}$$

其边界条件为

$$B(T,T) = 1 \tag{20.47}$$

求解该 PDE 可以得到封闭形式的定价公式，所得到的表达式比 Vasicek 情形更复杂，为

$$B(t,T) = A(t,T)e^{-C(t,T)r} \tag{20.48}$$

其中函数 $A(t,T)$、$C(t,T)$为

$$A(t,T) = \left(2\frac{\gamma e^{1/2(\alpha+\lambda+\gamma)T}}{(\alpha+\lambda+\gamma)(e^{\gamma T}-1)+2\gamma}\right)^{2\frac{\alpha\kappa}{b^2}} \tag{20.49}$$

$$C(t,T) = 2\frac{e^{\gamma^T}-1}{(\alpha+\lambda+\gamma)(e^{\gamma T}-1)+2\gamma} \tag{20.50}$$

γ 为

$$\gamma = \sqrt{(\alpha+\lambda)^2 + 2b^2} \tag{20.51}$$

我们可以采用相同的方法，得出在这种情形下的收益率曲线．

20.6 小结

本章讨论的是推导利率敏感型证券 PDE 的经典方法．我们可以看到，尽管主要步骤与 Black-Scholes 情形相似，两种情形在实践应用和基本思想方面还存在很大的差异．对利率 [342]

敏感型证券定价的经典方法依赖于对基础随机过程漂移项进行建模，而 Black-Scholes 方法只需对波动率建模和调整．

20.7 参考阅读

各种主流材料中都有关于债券价格的 PDE 解的叙述．但读者可能更倾向于 Vasicek 的原创性论文，论文在《Vasicek and Beyond》中能找到．另外两本很好的书是 Cox-Ingersoll-Ross(1985)及 Hull 和 White(1990).

20.8 习题

1. 假设瞬时即期利率的 SDE 为：

$$dr_t = \sigma r_t dW_t \tag{20.52}$$

其中 W_t 是在真实世界概率测度下的维纳过程，波动率 σ 为常数．起始即期利率 r_0 等于 5%.

(a) 该即期利率动态过程意味着什么？

(b) 在这些条件下计算无违约风险折价债券价格 $B(t,T)$的 PDE.

(c) 该 PDE 有解吗？

(d) 利率风险的市场价格等于多少？你能对其符号进行解释吗？

2. 已知即期利率模型：

$$dr_t = \alpha(\kappa - r_t)dt + bdW_t \tag{20.53}$$

其中 W_t 是真实世界概率测度下的维纳过程．根据该即期利率模型，相应无违约风险纯折价债券 $B(t,T)$PDE 的解给出了 $B(t,T)$的闭型债券定价公式：

$$B(t,T) = e^{\frac{1}{\alpha}(1-e^{-\alpha(T-t)})(R-r)-TR-\frac{b^2}{4\alpha^3}(1-e^{-\alpha(T-t)})^2} \tag{20.54}$$

其中

$$R = \kappa - \frac{b\lambda}{\alpha} - \frac{b^2}{\alpha^2} \tag{20.55}$$

对于这种折价债券，考虑下面的问题：

(a) 对上面的 $B(t,T)$应用伊藤引理，得到描述该债券动态过程的 SDE.

(b) 该债券动态过程的漂移项和扩散项是什么？推导出这些量确切的表达式，证明漂移系数 μ 等于：

$$\mu = r_t - \frac{b\lambda}{\alpha}(1 - e^{-\alpha(T-t)})$$

扩散系数等于：

$$\frac{b}{\alpha}(1 - e^{-\alpha(T-t)})$$

(c) 扩散系数与风险市场价格 λ 相互独立吗？

(d) 折价债券的期限与波动率之间有什么样的关系？

(e) 风险溢价(即超出无风险利率的收益率)与波动率成正比吗？与风险市场价格成正比吗？这一点重要吗？

(f) 当 $T \to \infty$时，漂移系数和扩散系数将发生什么变化？

(g) R 表示什么？

3. 短期利率 $r(t)$ 的标准双变量 Vasicek 模型为

$$dY_1(t) = -\lambda_1 Y_1(t)dt + dW_1(t) \tag{20.56}$$

343

$$dY_2(t) = -\lambda_{21} Y_1(t)dt - \lambda_2 Y_2(t)dt + dW_2(t) \tag{20.57}$$

$$r(t) = \delta_0 + \delta_1 Y_1(t) + \delta_2 Y_2(t) \tag{20.58}$$

其中 $W_1(t)$ 和 $W_2(t)$ 是在 Q 测度下相互独立的布朗运动．零息债券价格为

$$Z_t^T = E(e^{-\int_t^T r(s)ds}), \quad 0 \leqslant t \leqslant T \tag{20.59}$$

由于(20.1)和(20.2)的解是马尔可夫过程，必然存在函数 $f(t,y_1,y_2)$ 使得 $Z_t^T = f(t,Y_1(t),Y_2(t))$ 成立．证明 $f(t,y_1,y_2)$ 满足下面的 PDE：

$$-(\delta_0 + \delta_1 y_1 + \delta_2 y_2)f(t,y_1,y_2) + f_t(t,y_1,y_2) \tag{20.60}$$

$$-\lambda_1 y_1 f_{y_1}(t,y_1,y_2) - \lambda_{21} y_1 f_{y_2}(t,y_1,y_2) - \lambda_2 y_2 f_{y_2}(t,y_1,y_2) \tag{20.61}$$

$$+\frac{1}{2} f_{y_1 y_1}(t,y_1,y_2) + \frac{1}{2} f_{y_2 y_2}(t,y_1,y_2) = 0 \tag{20.62}$$

4. 考虑下面即期利率的 SDE：$dr_t = \sigma r_t dW_t$。

(a) 编写程序，通过模拟 dW_t 给出 r_t 的路径．假设 $r_0 = 0.01$，$T = 2.0$，$\Delta = 0.05$. 计算结束时即期利率 r_T 的期望值．

(b) 求解已知 SDE，并重复上面的拟合过程．同上面一样，计算结束时即期利率 r_T 的期望值．

(c) 证明这两种模拟方法在路径增加时是收敛的，并对它们相关有效性做出评价．

344

第21章 条件期望与 PDE 的联系

21.1 引言

纵观全书，我们一直在轮流讨论两种主要定价方法的数学工具．通过金融基本定理以及用现金账户标准化，我们得到了通常用于对衍生品定价的公式

$$F(S_t,t)=E_t^Q[\mathrm{e}^{-\int_t^T r_s\mathrm{d}s}F(S_T,T)] \tag{21.1}$$

其中该衍生品的标的资产为 S_t，到期日的收益为 $F(S_T,T)$．根据该公式，在无风险测度 Q 下未来收益用 $\mathrm{e}^{\int_t^T r_s\mathrm{d}s}$ 折现后的条件期望值等于当前无套利价格 $F(S_t,t)$．若 r_t 为常数，即像 Black-Scholes 假设一样，则公式可简化为：

345

$$F(S_t,t)=\mathrm{e}^{-r(T-t)}E_t^Q[F(S_T,T)] \tag{21.2}$$

在其他情况下，我们需要借助 PDE 进行定价．例如，在上一章，我们利用无风险投资组合的方法，得到了无套利条件下无违约折现债券 $B(t,T)$所满足的 PDE：

$$B_r(a(r_t,t)-\lambda_t b(r_t,t))+B_t+\frac{1}{2}B_{rr}b(r_t,T)^2-r_tB=0 \tag{21.3}$$

其边界条件为

$$B(T,T)=1 \tag{21.4}$$

同样，在 Black-Scholes 假设下，即期利率为常数 r，我们已经得到了看涨期权的 Black-Scholes PDE，其执行价格为 K，到期日为 T，标的证券价格为 S_t：

$$S_tF_sr_t+F_t+\frac{1}{2}F_{ss}\sigma_t^2-rF=0 \tag{21.5}$$

其边界条件为

$$F(S_T,T)=\max[S_T-K,0] \tag{21.6}$$

所以，定价工作主要使用 PDE 方法和条件期望方法．这两种方法最终得到的是同一个无套利价格 $F(S_t,t)$．这意味着可能在(21.1)和(21.2)那样的条件期望和(21.3)和(21.5)那样的 PDE 之间存在着深层联系．

实际上，假如我们证明对于如下函数 $F(S_t,t)$

$$F(S_t,t)=E_t^Q[\mathrm{e}^{-\int_t^T r_s\mathrm{d}s}F(S_T,T)] \tag{21.7}$$

若 $F(S_t,t)$是二阶可导的，则它将自动满足某个 PDE. 我们能够比较方便地得到这个 PDE 的一般形式．我们将讨论几个例子．

所有利率衍生品都必须假设瞬时即期利率是随机的．同时，由于金融基本定理，我们总可以将风险中性测度下的衍生品价格 $F(S_t,t)$写为

$$F(S_t,t)=E_t^Q[\mathrm{e}^{-\int_t^T r_s\mathrm{d}s}F(S_T,T)] \tag{21.8}$$

那么，我们就能得到衍生品价格的条件期望．对于利率衍生品尤其如此，此时即期利率不能假设为常数，而折现因子也必须是随机的．

但是这些条件期望常常难以计算．r_t 的随机性使得这项任务非常复杂．通常我们求不出封闭形式的解，只能求数值解．即使可以求出数值解，限于精确度和计算量，我们可能不得不采取别的方法．所以，我们十分有必要研究一种替代方法，来避免直接用条件期望计算无套利价格 $F(S_t,t)$. 特别地，如果我们能得到条件期望(21.1)或(21.2)对应的 PDE，我们就能用微分方程的数值方法计算 $F(S_t,t)$. 如果我们可以构造与这些期望对应的微分方程，我们就能用更加便捷、准确的数值方法来得到公平市场价格 $F(S_t,t)$. [⊖]

同样，如果市场参与者得到了一个难以求解的 PDE，若他能够得到类似(21.8)的期望形式，就可能能够更方便地求出 $F(S_t,t)$. 此时这种对应关系也非常实用． 346

本章我们将讨论这种对应关系的原理以及相关数学工具．

21.2 从条件期望到 PDE

本节我们将建立一类条件期望与 PDE 的对应关系．我们将利用简单的例子阐明，在某些非平凡条件下，我们可以利用某一类条件期望定义一个函数，我们就可以得到该函数所满足的 PDE. 其中一个主要的必要条件是该过程具有马尔可夫性质．

我们讨论的第一个例子对于市场参与者并不是直接有用的．但是能够帮助我们理解整个推导过程．而且，我们将逐步增加例子的复杂程度，并展示这些方法如何在实际的衍生品定价中运用．

21.2.1 例 1：常数贴现因子

考虑随机过程 $x_t \in [0, \infty)$ 的函数 $F(x_t)$，由以下条件期望定义：

$$F(x_t) = E_t^P\left[\int_t^{\infty} \mathrm{e}^{-\beta s} g(x_s)\mathrm{d}s\right] \tag{21.9}$$

其中 $\beta > 0$ 表示常数瞬时折现率，$g(\cdot)$ 为某个连续收益，依赖于随机过程 x_t 的值．$E_t^P[\cdot]$为在测度 P 下的期望，以信息集 I_t 为条件，两者当前都是未确定的．过程 x_t 服从 SDE：

$$\mathrm{d}x_t = \mu \mathrm{d}t + \sigma \mathrm{d}W_t \tag{21.10}$$

其中 μ，σ 为已知常数．

$F(x_t)$可以看作未来现金流 $g(x_s)$折现后的期望价值，$g(x_s)$依赖于 I_S 可测的随机变量 x_S. 折现因子 $\Delta > 0$ 是确定的．

显然，金融市场的现金流一般是用随机折现因子折现．但对于利率衍生品，我们暂时不考虑这一点．此时我们的目标是得到与(21.9)中的数学期望相对应的 PDE. 我们要逐步具体学习推导出 PDE 的步骤．一旦我们明白了这一过程，下面很容易就能引入随机的折现

⊖ 例如，在研究美式衍生品时，用数值方法解 PDE 往往更为方便，而不是用蒙特卡洛方法估算条件期望．

因子 .

现在我们逐步推导(21.9)对应的 PDE. 这些步骤是一般性的，可以应用到比(21.9)更复杂的情况中 . 我们用一种比较机械的方法推导并说明这一过程 . 为简化标记，我们假设初始时刻为 $t=0$.

首先，取时间区间为 $\Delta>0$ 并将$[0,\infty)$分为两份 . 一份是即近的未来，用$[0,\Delta)$表示，剩下的用$[\Delta,\infty)$表示，

$$F(x_0)=E_0^P\left[\int_0^\Delta e^{-\beta s}g(x_s)ds+\int_\Delta^\infty e^{-\beta s}g(x_s)ds\right] \tag{21.11}$$

第二步是运用初等变换，在表达式的右边引入 $F(\cdot)$的未来价值 . 通过将括号内的第二项乘以并除以 $e^{-\beta\Delta}$，我们可将其改写为：

347

$$E_0^P\left[\int_\Delta^\infty e^{-\beta s}g(x_s)ds\right]=E_0^P\left[e^{-\beta\Delta}\int_\Delta^\infty e^{-\beta(s-\Delta)}g(x_s)ds\right] \tag{21.12}$$

第三步将使用条件期望的递归性质 . 我们之前介绍过，若条件期望是嵌套的，只有依赖于较小信息集的期望有意义 . 所以，如果 $I_t\subseteq I_S$，则

$$E_t^P[E_s^P[\cdot]]=E_t^P[\cdot] \tag{21.13}$$

那么我们就可以用算子 $E_0^P[\cdot]$替代算子 $E_0^P[E_\Delta^P[\cdot]]$.[㊀]因此可得

$$E_0^P\left[\int_\Delta^\infty e^{-\beta s}g(x_s)ds\right]=E_0^P\left[e^{-\beta\Delta}E_\Delta^P\left[\int_\Delta^\infty e^{-\beta(s-\Delta)}g(x_s)ds\right]\right] \tag{21.14}$$

我们可以看到，右端内部括号中的项恰是 $F(x_\Delta)$,[㊁]即：

$$E_0^P\left[\int_\Delta^\infty e^{-\beta s}g(x_s)ds\right]=E_0^P[e^{-\beta\Delta}F(x_\Delta)] \tag{21.15}$$

可以将上式运用到(21.11)中，得：

$$F(x_0)=E_0^P\left[\int_0^\Delta e^{-\beta s}g(x_s)ds+e^{\beta\Delta}F(x_\Delta)\right] \tag{21.16}$$

将 $F(x_0)$移到右端的期望算子中，可得：

$$E_0^P\left[\int_0^\Delta e^{-\beta s}g(x_s)ds+e^{\beta\Delta}F(x_\Delta)-F(x_0)\right]=0 \tag{21.17}$$

第四步，先加后减去 $F(x_\Delta)$，再将所有项除以 Δ，重组后可得：

$$\frac{1}{\Delta}E_0^P\left[\int_0^\Delta e^{-\beta s}g(x_s)ds+(e^{\beta\Delta}-1)F(x_\Delta)+[F(x_\Delta)-F(x_0)]\right]=0 \tag{21.18}$$

最后一步，我们对左端关于 $\Delta\to0$ 取极限，第二项的括号内实际上是 $e^{\beta x}$ 的导数在 $x=0$ 处的值：

$$\lim_{\Delta\to0}\frac{1}{\Delta}(e^{\beta\Delta}-1)=-\beta \tag{21.19}$$

第一项是黎曼积分的上极限：

㊀ 在 $t=\Delta$ 时刻，我们拥有的信息比 $t=0$ 时刻更多 .

㊁ 这里 $F(x_0)$是 $F(\cdot)$在 $t=0$ 时刻的观测值 . 而 $F(x_\Delta)$是经过一小段时间 Δ 后在 $t=\Delta$ 的值 . 它将以 x_Δ 为条件 .

$$\lim_{\Delta \to 0} \frac{1}{\Delta} \int_0^{\Delta} e^{-\beta s} g(x_s) ds = g(x_0) \tag{21.20}$$

而第三项包含了随即积分的期望，因此需要使用伊藤引理．首先，我们用泰勒展开式近似，得：

$$\frac{1}{\Delta} E_0^P [F(x_\Delta) - F(x_0)] \approx \frac{1}{\Delta} E_0^P \left[F_x [x_\Delta - x_0] + \frac{1}{2} F_{xx} \sigma_x^2 \Delta \right] \tag{21.21}$$

然后令 $\Delta \to 0$ 并求期望，可得：

$$\lim_{\Delta \to 0} \frac{1}{\Delta} E_0^P \left[F_x [x_\Delta - x_0] + \frac{1}{2} F_{xx} \sigma_x^2 \Delta \right] = F_x \mu + \frac{1}{2} F_{xx} \sigma^2 \tag{21.22}$$

其中 μ 是随机过程 x_t 的漂移率，是对$(x_\Delta - x_0)$使用期望算子的结果．

将(21.19)～(21.22)带入(21.18)，即可得到我们想要的 PDE：

$$F_x \mu + \frac{1}{2} F_{xx} \sigma^2 - \beta F + g = 0 \tag{21.23}$$

其中 F_x、F_{xx}、F 以及 g 都是 x 的函数．

348

你可能好奇为什么条件期望(21.9)和该 PDE 之间存在联系．毕竟，这两个概念看起来毫无关系．下面是一个启发性的解释．

PDE 与现金流$\{g(x_S)\}$的“现值”的期望有关．如果我们用上述条件期望给出 $F(\cdot)$的现值，那么它必然是关于 x_0 满足一定条件的函数，且其动态过程必然会受到 x 未来变化期望的影响．这些限制条件导致了相应 PDE 的存在．

更具体地说，函数 $F(x_0)$是最优预测的结果．最优预测需要判断 $F(x_t)$可能随时间变化的方式．随机变量 x_t 的期望变化、时间变量 t 的确定性变化、收益 $g(x_t)$以及二阶伊藤校正项，这些都会对 $F(\cdot)$造成各种可预测的变化．最优预测应该将这些变化都考虑在内．对于所考虑的可预测变化，我们使预测误差的期望值为 0、方差最小化，从而得到与条件期望算子对应的 PDE. ⊖

21.2.2 例 2：债券定价

下面我们来看一个更贴近实际的例子．我们将用同样的方法，得到无违约风险纯折现债券价格的 PDE.

考虑无违约风险纯折现债券价格 $B(t,T)$，其到期日为 T 且无套利．假设瞬时即期利率 r_t 是马尔可夫过程，且债券的票面价值为 1 美元，利用我们熟悉的公式：

$$B(t,T) = E^Q [e^{-\int_t^T r_s ds}] \tag{21.24}$$

且

$$B(T,T) = 1$$

在求期望时，我们使用的是风险中性测度，且依赖于 t 时刻的信息，即 I_t．I_t 包含当前的

⊖ 注意在求 PDE 时，我们用 0 来代替 x_t 的维纳成分．

即期利率 r_t. 若 r_t 是马尔可夫过程，则 $B(t,T)$仅依赖于最近一次对 r_t 的观察．由于我们在风险中性世界中，根据 r_t 的含义，我们可以得出以下 SDE：

$$dr_t = [a(r_t,t) - \lambda_t b(r_t,t)]dt + b(r_t,t)dW_t \tag{21.25}$$

其中 W_t 是在风险中性测度 Q 下的维纳过程．λ_t 为利率风险市场价，具体定义为

$$\lambda_t = \frac{\mu - r_t}{\sigma} \tag{21.26}$$

其中 μ、σ 是债券公式中相应系数的简写：

$$dB = \mu(B,t)Bdt + \sigma(B,t)BdW_t$$

所以，和上一个例子一样，我们又得到了一组条件期望以及控制它的过程．这意味着我们可以运用同样的步骤得到 $B(t,T)$对应的 PDE. 而且在本例中，我们得到的 PDE 对于债券定价还有一些实际用途．我们可以求它的数值解，或者解析解(如果存在的话).

以上步骤可以直接套用，不必讨论细节．首先，将区间$[t,T]$分为两部分并写出：

349

$$B(t,T) = E^Q[(e^{-\int_t^{t+\Delta} r_s ds})(e^{-\int_{t+\Delta}^{T} r_s ds})] \tag{21.27}$$

第二步，尝试在该表达式中引入债券的未来价格 $B(t+\Delta, T)$. 事实上，如果我们运用条件期望的递归性质，可以发现右端的第二个指数项恰是 $B(t+\Delta, T)$. 利用

$$E_t^P[E_{t+\Delta}^P[\cdot]] = E_t^P[\cdot] \tag{21.28}$$

可得

$$B(t,T) = E^Q\left[\left(e^{-\int_t^{t+\Delta} r_s ds}\right)B(t+\Delta,T)\right] \tag{21.29}$$

因为

$$B(t+\Delta,T) = E^Q\left[e^{-\int_{t+\Delta}^{T} r_s ds}\right] \tag{21.30}$$

第三步，将所有项移到期望符号中，先增后减 $B(t+\Delta,T)$，然后除以 Δ：

$$\frac{1}{\Delta}E^Q\left[\left(e^{-\int_t^{t+\Delta} r_s ds} - 1\right)B(t+\Delta,T) + [B(t+\Delta,T) - B(t,T)]\right] = 0 \tag{21.31}$$

注意到左边含有增量 $B(t+\Delta,T) - B(t,T)$，我们可以通过它来运用伊藤引理．

第四步，关于 $\Delta\to 0$ 求极限，则第一项为：[⊖]

$$\lim_{\Delta\to 0}\frac{1}{\Delta}E^Q\left[\left(e^{-\int_t^{t+\Delta} r_s ds} - 1\right)B(t+\Delta,T)\right] = -r_t B(t,T) \tag{21.32}$$

然后对(21.31)的第二项运用伊藤引理，并求期望：

$$\lim_{\Delta\to 0}\frac{1}{\Delta}E^Q[B(t+\Delta,T) - B(t,T)]$$

$$= B_t + B_r[a(r_t,t) - \lambda_t b(r_t,t)] + \frac{1}{2}B_{rr}b(r_t,t)^2 \tag{21.33}$$

⊖ 这里我们假设满足相关的技术条件，允许极限和期望算子的互换．

其中 $a(r_t,t)$、$b(r_t,t)$分别是即期利率的漂移率和波动率．[1]

最后，将这些极限带入(21.31)，即可得到条件期望(21.24)对应的 PDE：

$$-r_tB+B_t+B_r[a(r_t,t)-\lambda_t b(r_t,t)]+\frac{1}{2}B_{rr}b(r_t,t)^2=0 \tag{21.34}$$

边界条件为

$$B(T,T)=1 \tag{21.35}$$

这就是无违约风险纯折现债券的无套利价格所满足的 PDE. 在第 20 章，我们曾用无风险投资组合的方法得到了该公式．

21.2.3 例 3：一般情况

我们刚刚介绍的两个例子，具体展现了如何从一类条件期望推导出对应的 PDE. 在例 1 中，现金流是随机的，依赖于标的过程 x_t，而折现率是常数．在例 2，债券在到期日支付一笔固定的现金流，但折现因子是随机的．

显然，我们可以将这两个基础例子结合起来，得到一种收益和折现率都随机的情况，进而求出 PDE. 我们设所研究的金融工具的收益为 $g(r_t)$，依赖于即期利率，而折现因子也是随机的：

$$F(r_t,t)=E_t^Q\left[\left(\mathrm{e}^{-\int_t^u r_s\mathrm{d}s}\right)g(r_u)\mathrm{d}u\right] \tag{21.36}$$

$F(\cdot)$代表一种收益与即期利率相关的工具的价格，期限为 $u\in[t,T]$，因此在每个 u 要采用随机折现因子 D_u：

350

$$D_u=\mathrm{e}^{-\int_t^u r_s\mathrm{d}s} \tag{21.37}$$

有趣的是，D_u 的期望正是 u 时刻支付 1 美元的无违约风险纯折现债券在 t 时刻的价格．[2]

许多金融工具和利率衍生品都属于这类，如折现债券，金融期货等，它们的无套利价格可由类似(21.36)的期望形式给出．所以，只要驱动这些期望的过程是马尔可夫的，我们就可以用上面两节的方法找到对应的 PDE. 而这些 PDE 可是在实际的金融市场中帮助我们研究上述的复杂的金融工具．

21.2.4 一些说明

我们需要说明一些内容，以免读者难以理解．

马尔可夫条件的重要性

在我们上面研究的这一类期望中，驱动它们变化的随机过程都是满足马尔可夫条件

---

[1] 和上一个例子不同，这里函数 $B(t,T)$依赖于 t 和 r_t. 所以，存在一个之前没有的 B_t 项．

[2] 这里我们不能直接对 D_u 使用算子，因为 $g(r_u)$与 D_u 相关．如果这种相关性不存在，且 $g(\cdot)$仅依赖于一个独立的随机变量，如 x_u，那么我们就可以分别求期望，然后用对应的折现价格 B_u 乘以收益，从而折现：

$$E_t^Q\left[\int_t^T \mathrm{e}^{-\int_t^T r_s\mathrm{d}s}g(x_u)\mathrm{d}u\right]=\int_t^T B_uE_t^Q[g(x_u)]\mathrm{d}u \tag{21.38}$$

假设算子的互换是允许的．另一方面，如果我们使用如第 17 章讨论的远期测度，方程(21.38)总成立．

的. 只有满足该条件，以上推导才成立. 所以我们有必要弄清楚马尔可夫假设在推导中起了什么作用.

在推导 PDE 时，我们使用了条件期望算子 $E^Q[\cdot]$，现在我们写出它的展开形式，即具体写出它的条件信息集：

$$E^Q\left[\int_t^T \mathrm{e}^{-\int_t^T r_s \mathrm{d}s} g(x_u)\mathrm{d}u \middle| I_t\right] = E_t^Q\left[\int_t^T B_u[g(x_u)]\mathrm{d}u \middle| r_t\right] \tag{21.39}$$

$$= F(r_t, t) \tag{21.40}$$

只有当 r_t 是马尔可夫过程时以上运算才成立. 如果没有该假设，那么条件期望还将依赖于除了 r_t 以外的因素. 也就是说，过去的即期利率$\{r_s, s<t\}$也将成为价格的决定因素. 同时，该金融工具的价格不能再记作 $F(t, r_t)$，因为这表示它只依赖于 t 和 r_t. 其余的推导也都不再成立.

所以，如果要用上面的方法研究利率衍生品，那么标的过程必须满足马尔可夫条件.

21.2.5 哪一种漂移率

读者可能想知道在以上 PDE 推导中我们应该取什么参数作为随机过程的漂移率. 结论我们之前已经给出了，但是还有必要进行说明.

我们所研究的条件期望是在某个(条件)概率分布下给出的. 例如，如果我们将债券的无套利价格表示为：

$$B(t, T) = E_t^Q[\mathrm{e}^{-\int_t^T r_s \mathrm{d}s}] \tag{21.41}$$

那么我们是关于风险中性测度 Q 求期望. 对于所研究的随机过程 r_t，由于我们选择的是风险中性测度，我们也需要对 r_t 取风险调整漂移率，对应的 SDE 为：

351

$$\mathrm{d}r_t = (a(r_t, t) - \lambda_t b(r_t, t))\mathrm{d}t + b(r_t, t)\mathrm{d}W_t \tag{21.42}$$

而不是“现实世界”SDE：

$$\mathrm{d}r_t = a(r_t, t)\mathrm{d}t + b(r_t, t)\mathrm{d}W_t^* \tag{21.43}$$

其中 W_t^* 为在现实世界概率 P 下的维纳过程.

所以，在当前设定下使用伊藤引理时，如果要用到漂移率，我们必须使用$(a(r_t, t) - \lambda_t b(r_t, t))$而非 $a(r_t, t)$. 在(21.33)中取极限时就是如此.

那是否会用到非调整的漂移率呢? 这个问题很有意思，因为这样做意味着我们不会使用风险中性测度的定价方法. 最终的公式本质上应该是一样的，但是可能不太实用. 也就是说，这个问题恰恰体现出了鞅方法的作用.

确实，在推导过程中，我们可以不用风险调整漂移率，而是用即期利率过程的原始漂移率. 但这样我们必须在真实概率 P 下求条件期望，而非风险中性概率. 但是我们知道表达式诸如

$$B(t, T) = E_t^P[\mathrm{e}^{-\int_t^T r_s \mathrm{d}s}] \tag{21.44}$$

一般不能成立，只要 $B(t, T)$无套利且使用真实概率 P. 如果一定要使用真实概率，那么无

套利价格的公式为：

$$B(t,T) = E_t^P\left[e^{-\int_t^T r_s ds} e^{\int_t^T [\lambda_s(r_s,s) dW_t^* - \frac{1}{2}\lambda(r_s,s)^2 ds]}\right] \tag{21.45}$$

其中所有符号含义与(21.42)和(21.43)相同．

通过将上述条件期望拆分然后使用和之前一样的步骤，我们还是能得到(21.34)中的PDE. 最主要的区别存在于(21.33)对应的极限计算中．此时我们必须代入真实世界的漂移率$a(r_t,t)$而非风险调整漂移率．

21.2.6 另一个债券价格公式

本章的主要重点是 PDE 和条件期望的联系．不过，我们还要顺带地介绍等价鞅测度在债券定价中的一类应用．

前面介绍了两个债券定价公式．一种使用鞅测度得到表达式：

$$B(t,T) = E_t^Q\left[e^{-\int_t^T r_s ds}\right] \tag{21.46}$$

另一种使用真实概率 P 得到

$$B(t,T) = E_t^P\left[e^{-\int_t^T r_s ds} e^{\int_t^T [\lambda(r_s,s) dW_u^* - \frac{1}{2}\lambda(r_s,s)^2 ds]}\right] \tag{21.47}$$

当然，这两个 $B(t,T)$是完全相同的，只是定义和计算的方式不同．

现在我们研究的问题是如何从其中一个公式推算出另一个．这是一个关于 Girsanov 定理应用的很好的例子．首先，我们在第 15 章中介绍过，概率 P 和 Q 是等价的如果存在以下关联：

$$dQ_t = \xi_t dP_t \tag{21.48}$$

其中 Radon-Nikodym 导数 ξ_t 为

$$\xi_t = e^{\int_0^t [\lambda_u dW_u^* - \frac{1}{2}\lambda_u^2 du]} \tag{21.49}$$

352

其中 ξ_t 是 I_t 可测过程．[⊖]

假定所有 Girsanov 定理的前提条件都满足，我们现在介绍如何根据(21.46)得到(21.47).

从债券定价方程开始：

$$B(t,T) = E_t^Q\left[e^{-\int_t^T r_s ds}\right] \tag{21.50}$$

利用条件期望算子 $E_t^Q[\cdot]$的定义写出该表达式：

$$E_t^Q\left[e^{-\int_t^T r_s ds}\right] = \int_\Omega \left(e^{-\int_t^T r_s ds}\right) dQ \tag{21.51}$$

其中 Ω 是未来 r_t 将要取值的范围．现在利用(21.48)中 P 和 Q 的等价关系替换方程中的 dQ：

⊖ 在这个例子中，λ_t 是即期利率风险的市场价格．

$$E_t^Q\left[\mathrm{e}^{-\int_t^T r_s \mathrm{d}s}\right] = \int_\Omega \left(\mathrm{e}^{-\int_t^T r_s \mathrm{d}s}\right)\xi_T \mathrm{d}P \tag{21.52}$$

替换 ξ_T，即可得到目标方程：

$$E_t^Q[\mathrm{e}^{-\int_t^T r_s \mathrm{d}s}] = \int_\Omega \left(\mathrm{e}^{-\int_t^T r_s \mathrm{d}s}\mathrm{e}^{\int_0^T[\lambda_u \mathrm{d}W_u^* - \frac{1}{2}\lambda_u^2 \mathrm{d}u]}\right)\mathrm{d}P \tag{21.53}$$

$$= E_t^P\left[\mathrm{e}^{-\int_t^T r_s \mathrm{d}s}\mathrm{e}^{\int_t^T[\lambda(r_s,s)\mathrm{d}W_u^* - \frac{1}{2}\lambda(r_s,s)^2 \mathrm{d}s]}\right] \tag{21.54}$$

这就是我们之前得到的真实概率下的债券定价公式．

所以，一旦使用 Girsanov 定理，两种无违约风险纯折现债券价格公式之间的联系就变得非常简单．当然，在上面的推导中，我们没有说明 λ_t 是利率风险的市场价格，但它很明显是对利率随机微分方程漂移率的调整．

21.2.7 用哪一个公式

表达式(21.46)和(21.54)分别给出了两种不同的描述 $B(t,T)$的方法．其中，由真实概率测度所推导出的公式(21.54)中包含 λ_t，而(21.46)却不含这一项，这使得(21.54)在实际计算中可能会更复杂．那么当人们在使用蒙特卡洛方法估计债券价格时，可能会倾向于使用(21.46)，以为这样可以避开关于 λ_t 的估计或者计算．

可是，这只是表象而已．实际上，不管使用这两种方法中的哪一种，λ_t 的估算在蒙特卡洛定价中都是必需的．在(21.54)中，这一点显然：因为公式本身就自然包含 λ_t. 而在(21.46)中，我们一样需要对 λ_t 做出相应的数值估计．有了这些估计，我们才可以通过下面这个在鞅测度下的 SDE 去模拟 r_t 的路径：

$$\mathrm{d}r_t = (a(r_t,t) - \lambda_t b(r_t,t))\mathrm{d}t + b(r_t,t)\mathrm{d}W_t \tag{21.55}$$

显然，只有在知道了 λ_t 后，这个公式才能被有效地使用．

综上所述，我们可以发现，在(21.54)中，最后的积分式中含有 λ_t，但是对应的 SDE 中不含 λ_t；而在(21.46)正相反，其所对应的 SDE 中含有 λ_t，而积分中不含这一项．对于蒙特卡洛方法而言，SDE 和最后的期望积分都会被用到．这就是为什么我们到现在为止所讨论的方法都只是“经典的”方法：不管如何，我们还是必须对漂移项进行建模．而我们在后面所介绍的 HJM 方法则可以避开这一困难．

21.3 从 PDE 到条件期望

到目前我们已经说明，如果标的过程是马尔可夫的，且满足某些数学意义的条件，那么在某个合适测度下的条件期望定义的无套利价格必定满足某个 PDE. 也就是说，已知一个满足一定条件的条件期望，我们就能得到它对应的 PDE.

353

本节我们的研究与之前方向相反的问题．假设我们已知某资产价格 $F(S_t,t)$所满足的 PDE，我们能否得到对应的条件期望呢?

我们通过一个特殊情况讨论该问题．我们令 $F(W_t,t)$为依赖于在某概率测度下的维纳过程 W_t 的金融衍生品价格．选择 W_t 作为驱动过程可能不是非常贴合实际，但它很容易被一般化．此外，它可以套用到工程学中的热传导方程上．

假设衍生品价格 $F(W_t,t)$满足以下 PDE：

$$F_t + \frac{1}{2}F_{ww} = 0 \tag{21.56}$$

同时我们有以下 $t=T$ 处的边界条件：

$$F(W_T,T) = G(W_T)$$

$G(\cdot)$为某个已知函数．

下面我们证明这个 PDE 的解可以表示为一个条件期望．为此，我们首先假设所有技术性条件都满足，然后对 $F(W_t,t)$使用伊藤引理：

$$\mathrm{d}F = \left[\frac{\partial F}{\partial t} + \frac{1}{2}\frac{\partial^2 F}{\partial W^2}\right]\mathrm{d}t + \frac{\partial F}{\partial W}\mathrm{d}W_t \tag{21.57}$$

$$= \left[F_t + \frac{1}{2}F_{ww}\right]\mathrm{d}t + F_w\mathrm{d}W_t \tag{21.58}$$

其中我们利用了维纳过程的性质，即漂移参数等于 0，扩散参数等于 1.

这个随机微分方程展现了 $F(W_t,\ t)$如何随时间变化．下一步是对两边从 t 到 T 积分：

$$F(W_T,T) - F(W_t,t) = \int_t^T F_w\mathrm{d}W_s + \int_t^T\left[F_t + \frac{1}{2}F_{ww}\right]\mathrm{d}s \tag{21.59}$$

不要忘记偏导数 F_S 和 F_{WW}也是 W_S 和 S 的函数．

我们可以求右端的积分．事实上，根据(21.56)的 PDE，第二个积分项等于 0：

$$\int_t^T\left[F_t + \frac{1}{2}F_{ww}\right]\mathrm{d}s \tag{21.60}$$

进一步，对(21.59)两端求期望，可得

$$E_t^Q[F(W_T,T)] = F(W_t,t) + E_t^Q\left[\int_t^T F_w\mathrm{d}W_s\right] \tag{21.61}$$

$F(W_T,T)$是 $F(\cdot)$在边界 $t=T$ 的值，所以我们可以用已知的 $G(W_T)$替换．整理后可得：

$$F(W_t,t) = E_t^Q[G(W_T)] - E_t^Q\left[\int_t^T F_w\mathrm{d}W_s\right] \tag{21.62}$$

所以，如果我们能证明右端第二个期望为 0，那么(未知)函数 $F(\cdot)$即可用已知函数$G(\cdot)$的期望来定义．但这要求：

$$E_t^Q\left[\int_t^T \frac{\partial F}{\partial W}\mathrm{d}W_s\right] = 0 \tag{21.63}$$

为证明该式成立，我们要用到关于维纳过程的伊藤积分的一个重要性质．从第 10 章我们知道，如果 $h(W_t)$是关于信息集 I_t 在测度 P 下的非预报函数，那么它关于 W_t 积分的期望值将为 0：

$$E_t^Q\left[\int_t^T h(W_s)\mathrm{d}W_s\right] = 0 \tag{21.64}$$

354

我们回顾一下原因 . W_t 是维纳过程 . 其增量 $\mathrm{d}W_t$ 不依赖于过去，包括刚过去的瞬间 . 但是如果 $h(W_t)$是非预测的，那么 $h(W_t)$也不依赖于“将来”. 所以，在(21.57)中需要相乘的项都是相互独立的，而其中之一即 $\mathrm{d}W_t$ 的均值为 0.

回到方程(21.63)，我们要使该项等于 0，而该项

$$E_t^Q\left[\int_t^T \frac{\partial F}{\partial W}\mathrm{d}W_s\right] \tag{21.65}$$

正是我们刚才介绍的那种类型 . 它是非预测函数关于维纳过程的积分 . 这意味着只要 $F(\cdot)$满足一定条件，则它的期望为 0：

$$E_t^Q\left[\int_t^T \frac{\partial F}{\partial W}\mathrm{d}W_s\right]=0 \tag{21.66}$$

所以我们有

$$F(W_t,t) = E_t^Q[G(W_T)] \tag{21.67}$$

这是价格 $F(W_t,\ t)$的另一种表达方法 . 即以 $G(W_T)$为边界条件、在概率 Q 下的条件期望形式. 该函数还是热传导方程的解 . 事实上，从一个包含未知函数 $F(t,W_t)$的 PDE 出发，我们用已知函数在某概率下的期望的形式来定义解，其中含有维纳过程 W_t.

21.4 生成元、Feynman-Kac 公式和其他工具

根据之前所介绍的内容，我们不会对随机过程理论可以系统性地解决类似问题这一点再感到惊讶 . 这些数学工具都可以使标记更为简化、推导更为机械化 . 生成元这一概念正是如此，它等价于得到类似(21.33)的极限 . 此外还有 Feynman-Kac 定理，它为一类 PDE 给出了概率形式的解 . 我们前面的讨论其实已经涵盖了这些内容，下面我们将给出这些概念的正确叙述，作为本章的结尾 .

21.4.1 伊藤扩散过程

设连续随机过程 S_t 具有有限的一阶矩和二阶矩，它满足以下一般形式的 SDE：

$$\mathrm{d}S_t = a(S_t,t)\mathrm{d}t+\sigma(S_t,t)\mathrm{d}W_t,\quad t\in[0,\infty) \tag{21.68}$$

我们现在假设漂移和扩散参数只依赖于 S_t[⊖]，则 SDE 可记为：

$$\mathrm{d}S_t = a(S_t)\mathrm{d}t+\sigma(S_t)\mathrm{d}W_t,\quad t\in[0,\infty) \tag{21.69}$$

其中 $a(\cdot)$和 $\sigma(\cdot)$分别为漂移参数和扩散参数 . 具有这种特征的过程叫做时齐的伊藤扩散过程 . 这类过程的瞬时漂移和扩散参数不直接依赖于 t. $a(\cdot)$和 $\sigma(\cdot)$满足适当的条件，这使得它们不会变化过于快 .

我们下面介绍两种伊藤扩散过程的性质 .

⊖ 在几乎所有收益无跳跃的情况中，实际中使用的 SDE 要么是几何型的，要么是均值回归型的 . 后者在利率衍生品种特别流行，因为人们广泛接受短期利率具有均值回归的特征 . 在这些条件下，漂移和扩散参数只是 S_t 的函数 . 而对时间的依赖性可以用来匹配原始期限结构 .

21.4.2 马尔可夫性质

该性质的内容我们之前已经介绍过．设 S_t 为伊藤扩散过程且满足 SDE：

$$dS_t = a(S_t)dt + \sigma(S_t)dW_t, \quad t \in [0,\infty) \tag{21.70}$$

任取一个有界函数 $f(\cdot)$，同时假设信息集 I_t 包含所有 t 之前的 S_u，$u\leqslant t$. 那么我们说 S_t 具有马尔可夫性质，只要：

$$E[f(S_{t+h}) \mid I_t] = E[f(S_{t+h}) \mid S_t], \quad h > 0 \tag{21.71}$$

也就是说，如果要预测 S_t 的未来变化，根据 t 以及之前的观察数据进行预测，效果等同于将该过程视为从 t 时刻开始的新过程．换句话说，在已知 S_t 的情况下，过去的信息不能帮助我们更好地预测．

21.4.3 伊藤扩散过程的生成元

设 S_t 为(21.70)中的伊藤扩散过程．设 $f(S_t)$为关于 S_t 二次可微的函数，S_t 在 t 时刻的值为 s_t.

下面研究从当前状态 s_t 开始，$f(S_t)$会如何变化．为此我们定义一个算子来表示这种变化．我们将 A 定义为 $f(S_t)$的期望变化率：

$$Af(s_t) = \lim_{\Delta\to 0} \frac{E[f(S_{t+\Delta}) \mid f(s_t)] - f(s_t)}{\Delta} \tag{21.72}$$

这里，小写的 s_t 表示 S_t 的一个观察值．右端表达式的分子衡量了 $f(S_t)$的期望变化．当我们除以 Δ 后，算子 A 改为测量变化率．在随机过程理论中，A 叫做伊藤扩散过程 S_t 的生成元．

一些读者可能想知道我们该如何定义 $f(S_t)$的变化率，因为它实际上是维纳过程的函数. 变化率就像导数一样，但我们已经证明过维纳过程是不可导的．所以可能有人会问，我们该如何论证 A 的存在性呢?

该问题的答案很简单．A 并不是要求出 $f(S_t)$的真实变化率，而是表示期望变化率．虽然维纳过程波动剧烈且不可导，但 $f(S_t)$的期望变化是更平滑的函数，在一定条件下，可以定义上述极限．[⊖]

21.4.4 A 的表示方法

首先注意到 A 是用极限表示的期望变化率．也就是说，我们用无穷小的时间变化考虑瞬时的将来．那么很明显，这样的变化与伊藤引理有关．事实上，在当前情况下当 S_t 是单变量随机过程时：

$$dS_t = a(S_t)dt + \sigma(S_t)dW_t, \quad t \in [0,\infty) \tag{21.73}$$

算子 A 为：

⊖ 每个期望代表一个平均．根据定义，平均值的变化比具体值更为平滑．

$$Af = a_t \frac{\partial f}{\partial S} + \frac{1}{2}\sigma_t^2 \frac{\partial^2 f}{\partial S^2} \tag{21.74}$$

下面我们将它与伊藤引理的结果进行比较．对 $f(S_t)$ 使用伊藤引理（S_t 由(21.73)给出），则有：

$$\mathrm{d}f(S_t) = \left[a_t \frac{\partial f}{\partial S} + \frac{1}{2}\sigma_t^2 \frac{\partial^2 f}{\partial S^2}\right]\mathrm{d}t + \sigma_t \frac{\partial f}{\partial S}\mathrm{d}W_t \tag{21.75}$$

所以，算子 A 和伊藤引理的结果主要有两点不同：

1)伊藤公式中的 $\mathrm{d}W_t$ 项被它的漂移项代替，也就是 0.

2)伊藤公式中的其余项被除以了 $\mathrm{d}t$.

356

这两点区别与 A 的定义是相容的．正如上面所提到的，A 计算的是从瞬时状态 s_t 开始的期望变化率．

多变量的情况

为了内容的完整性，我们还要介绍 A 的多变量情况．

设 X_t 为 k 维伊藤扩散过程，服从以下(向量)SDE：

$$\begin{bmatrix}\mathrm{d}X_{t_1}\\ \vdots \\ \mathrm{d}X_{t_k}\end{bmatrix} = \begin{bmatrix}a_{1,t}\\ \vdots \\ a_{k,t}\end{bmatrix}\mathrm{d}t + \begin{bmatrix}\sigma_t^{11} & \cdots & \sigma_t^{1k}\\ \vdots & & \vdots \\ \sigma_t^{k1} & \cdots & \sigma_t^{kk}\end{bmatrix}\begin{bmatrix}\mathrm{d}W_{1t}\\ \vdots \\ \mathrm{d}W_{1t}\end{bmatrix} \tag{21.76}$$

其中 a_{it} 是漂移系数，依赖于 t；σ_{it} 是扩散系数，也可能依赖于 t. 该方程可以写作以下符号形式：

$$\mathrm{d}X_t = a_t\mathrm{d}t + \sigma_t\mathrm{d}W_t \tag{21.77}$$

其中 $a(\cdot)$是 $k\times1$ 的向量，σ_t 是 $k\times k$ 的矩阵．

对应的算子 A 变化为：

$$Af = \sum_{i=1}^{k} a_{it} \frac{\partial f}{\partial X_i} + \sum_{i=1}^{k}\sum_{j=1}^{k} \frac{1}{2}(\sigma\sigma^{\mathrm{T}})^{ij} \frac{\partial^2 f}{\partial X_i\,\partial X_j} \tag{21.78}$$

其中$(\sigma\sigma^{\mathrm{T}})^{ij}$项代表矩阵$(\sigma\sigma^{\mathrm{T}})$的第 i 行第 j 列元素．

在大多数随机微积分的高阶著作中，所介绍的 A 的多变量形式都与上式相同．(21.78)中的表达式被称作 $f(\cdot)$的无穷小生成元．

21.4.5 Kolmogorov 向后方程

假设我们已知伊藤漂移过程 S_t，同时我们有 S_t 的函数 $f(S_t)$. 考虑以下期望：

$$\hat{f}(S^-,t) = E[f(S_t) \mid S^-] \tag{21.79}$$

其中 $\hat{f}(S^-,t)$表示预测值，S^- 表示 t 时刻以前最近一次的观测值．可以这样理解，S^- 是瞬时过去．那么，利用算子 A，我们可以推测 $\hat{f}(S^-,t)$随时间变化的特征．这种预测可以通过 Kolmogorov 后向方程得到：

$$\frac{\partial \hat{f}}{\partial t} = A\hat{f} \tag{21.80}$$

还记得 A 的定义为：

$$A\hat{f} = a_t \frac{\partial \hat{f}}{\partial S} + \frac{1}{2}\sigma^2 \frac{\partial^2 \hat{f}}{\partial S^2} \tag{21.81}$$

容易发现(21.81)中的等式正是 PDE：

$$\hat{f}_t = a_t \hat{f}_s + \frac{1}{2}\sigma_t^2 \hat{f}_{ss} \tag{21.82}$$

所以，我们又一次看到条件期望如

$$\hat{f}(S^-, t) = E[f(S_t) \mid S^-] \tag{21.83}$$

与(21.81)中的 PDE 的重要联系．和之前一样，这种对应关系可以用两种不同的方法描述：

- 期望(21.83)满足 PDE(21.81).
- 已知 PDE(21.81)，我们可以找到一个 $\hat{f}$ 满足该 PDE.

这样的结果意味着 $\hat{f}(S^-, t)$是(21.81)中 PDE 的解．所以，Kolmogorov 向后方程是随机过程的期望与 PDE 之间存在关联的例子，这就和本章之前介绍的一样．

例子

考虑以下函数：

$$p(S_t, S_0, t) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(S_t - S_0)^2}{2t}} \tag{21.84}$$

观察可以发现，这个函数是维纳过程的条件密度函数，该维纳过程在 $t=0$ 时刻从 S_0 出发，漂移率为 0，方差为 t. 357

如果我们能够写出该过程的随机微分方程，我们必须取漂移参数为 0，扩散参数为 1. 则 dS_t 满足：

$$dS_t = dW_t \tag{21.85}$$

我们对该密度函数使用 Kolmogorov 公式．由于该密度函数关于 S_t 二阶可导，它服从 Kolmogorov 向后方程

$$\hat{f}_t = a_t \hat{f}_s + \frac{1}{2}\sigma_t^2 \hat{f}_{ss} \tag{21.86}$$

但是根据(21.85)，在本例中，我们有：

$$a_t = 0 \tag{21.87}$$

以及

$$\sigma_t^2 = 1 \tag{21.88}$$

将这些参数带入，Kolmogorov 向后方程变为：

$$\hat{f}_t = \frac{1}{2}\hat{f}_{ss} \tag{21.89}$$

结果显示条件密度 $p(S_t, S_0, t)$是这样的一个函数．也就是说，对它关于 t 求一阶导、关于 S 求二阶导，带入上式，等式将成立．

根据这个结果，一个(一般化的)维纳过程的条件密度函数满足 Kolmogorov 向后方程．

该 PDE 告诉我们概率与 S_t 随时间的变化的联系．

21.5 Feynman-Kac 公式

Feynman-Kac 公式是 Kolmogorov 向后方程的一个拓展，我们在之前章节的讨论中提到过，下面给出正式的介绍．在已知 PDE 的情况下，通过该公式可以得到对应的概率解 $\hat{f}$．

Feynman-Kac 公式：已知

$$\hat{f}(t,r_t) = E^P[\mathrm{e}^{-\int_t^u q(r_s)\mathrm{d}s} f(r_u)] \tag{21.90}$$

我们有

$$\frac{\partial \hat{f}}{\partial t} = A\hat{f} - q(r_t)\hat{f} \tag{21.91}$$

其中算子 A 为

$$A\hat{f} = a_t \frac{\partial \hat{f}}{\partial r_t} + \frac{1}{2}\sigma_t^2 \frac{\partial^2 \hat{f}}{\partial r_t^2} \tag{21.92}$$

因此，对于特定的一类 PDE，Feynman-Kac 公式给出了条件期望形式的解．

21.6 小结

某些条件期望和 PDE 的联系在实际资产定价中非常实用．若已知一种具有特殊性质的金融工具，市场参与者可以利用这种联系得到隐含的 PDE，进而求数值解．

21.7 参考阅读

读者可以从 Kushner(2000)中找到一些关于关联性的有趣例子．这本书还给出了一些计算隐含 PDE 的实用方法．

21.8 习题

1. 假设债券价格 $B(t,T)$满足一下 PDE：

$$-r_t B + B_t + B_r(\mu - \lambda\sigma^B) + \frac{1}{2}B_{rr}\sigma^2 = 0 \tag{21.93}$$

358

$$B(T,T) = 1 \tag{21.94}$$

定义变量 $V(u)$：

$$V(u) = \mathrm{e}^{-\int_t^u r_s \mathrm{d}s} \mathrm{e}^{\int_t^u [\lambda(r_s,s)W_s - \frac{1}{2}\lambda(r_s,s)^2 \mathrm{d}s]} \tag{21.95}$$

其中 λ_s 是利率风险市场价格．

(a) $B(t,T)$为债券价格．计算 $\mathrm{d}(BV)$.

(b) 利用(21.93)中的 PDE 得到 $\mathrm{d}B(t,T)$的表达式．

(c) 对该表达式从 t 到 T 积分，并关于鞅等式求期望，得到债券定价公式：

$$B(t,T) = E_t^P[\mathrm{e}^{-\int_t^T r_s \mathrm{d}s} \mathrm{e}^{\int_t^T [\lambda(r_s,s)\mathrm{d}W_s - \frac{1}{2}\lambda(r_s,s)^2 \mathrm{d}s]}] \tag{21.96}$$

其中期望是以已知的当前利率 r_t 为条件的．

2. 记 r_t 为 t 时刻的短期利率，假设我们有仿射期限结构模型

$$Z_t^T = e^{A(t,T)+B(t,T)r_t} \tag{21.97}$$

其中 $A(t,T)$和 $B(t,T)$是确定型函数，t 为当前时刻，T 为到期日．

(a) 记 $f(t;s,s+\delta)$为 t 时刻的远期利率，期限为从 s 到 $s+\delta$. 计算 $f(t;s,s+\delta)$关于 r_t 和函数 $A(.,.)$以及 $B(.,.)$的表达式．

(b) 记 $r_s(s+\delta)$为 s 时刻的即期利率，借贷的到期日为 $s+\delta$. 给出 $r_s(s+\delta)$关于 r_t 和函数 $A(.,.)$以及 $B(.,.)$的表达式．

3. 考虑以下即期利率动态变化所满足的方程：

$$dr_t = \alpha(\kappa - r_t)dt + \sigma r_t^{\beta} dW_t \tag{21.98}$$

编写模拟程序计算条件期望：

$$P = E(e^{-\int_0^T r_s ds} \max(r_T - K, 0)) \tag{21.99}$$

取 $K=0.05$，$r_0=0.052$，$T=1$，$\Delta=0.01$. 此外，取 $\alpha=0.9$，$\kappa=0.07$，$\sigma=0.2$，$\beta=0.75$.

359

第 22 章　用傅里叶变换进行衍生品定价

在衍生品定价的一开始，我们假设标的资产的变化服从某个随机过程，其中标的资产可以是股票、商品、利率或者汇率．在利用条件密度函数对欧式看涨/看跌期权定价时，我们用折现收益关于条件密度求积分．不失一般性，我们仅以看涨期权为例．我们先介绍以下定义，令：

S_0 是标的证券今天的价格，

S_T 是标的证券 T 时刻的价格，

$f(S_T)\equiv f(S_T \mid S_0)$是 S_T 在风险中性测度下的条件概率密度函数，

$q(s_T)\equiv q(s_T \mid s_0)$是标的证券求对数后的密度函数，$s_T=\ln(S_T)$，

$k=\ln(K)$是执行价格的对数，

$C(K)\equiv\hat{C}(K,S_0,T)$是 T 时刻到期、执行价格为 K 的看涨期权价格，

$\hat{C}(k)$是 T 时刻到期、执行价格的对数为 k，$k=\ln(K)$的看涨期权价格．

假设无风险利率为 r，则期权价格可表示为

$$C(K)=\int_0^{\infty}\mathrm{e}^{-rT}(S_T-K)^+ f(S_T \mid S_0)\mathrm{d}S_T \tag{22.1}$$

$$=\int_K^{\infty}\mathrm{e}^{-rT}(S_T-K)f(S_T)\mathrm{d}S_T \tag{22.2}$$

通过变量代换 $s=\ln S$，$k=\ln K$，我们有

$$\hat{C}(k)=\int_k^{\infty}\mathrm{e}^{-rT}(\mathrm{e}^{s_T}-\mathrm{e}^{k})q(s_T \mid s_0)\mathrm{d}s_T \tag{22.3}$$

$$=\mathrm{e}^{-rT}\int_k^{\infty}(\mathrm{e}^{s_T}-\mathrm{e}^{k})q(s_T)\mathrm{d}s_T \tag{22.4}$$

在定价时，我们要用数值方法计算这种积分．我们可以使用各种数值积分方法，如梯形公式或者辛普森公式．已知当 s_T 趋于无穷时 $q(s_T|s_0)$趋于 0，我们可以取上界 $b=\ln B$ 并将区间 n 等分，格点为 $x_j=k+j\Delta x$，$j=0$，1，…，n，其中 $\Delta x=\dfrac{b-k}{n}$，

$$\hat{C}(k)=\mathrm{e}^{-rT}\int_k^{\infty}(\mathrm{e}^{s}-\mathrm{e}^{k})q(s)\mathrm{d}s \tag{22.5}$$

$$\approx\mathrm{e}^{-rT}\int_k^{b}(\mathrm{e}^{s}-\mathrm{e}^{k})q(s)\mathrm{d}s \tag{22.6}$$

$$=\mathrm{e}^{-rT}\sum_{i=0}^{n-1}\int_{x_j}^{x_{j+1}}(\mathrm{e}^{s}-\mathrm{e}^{k})q(s)\mathrm{d}s \tag{22.7}$$

当使用梯形公式时，我们用经过小区间端点的直线替代函数$(\mathrm{e}^{s}-\mathrm{e}^{k})q(s)$，然后进行积分．这相当于一开始用梯形面积来近似 $\int_{x_j}^{x_{j+1}}(\mathrm{e}^{s}-\mathrm{e}^{k})q(s)\mathrm{d}s$，即

$$\int_{x_j}^{x_{j+1}}(\mathrm{e}^s-\mathrm{e}^k)q(s)\mathrm{d}s \approx \frac{x_{j+1}-x_j}{2}[(\mathrm{e}^{x_j+1}-\mathrm{e}^k)q(x_{j+1})+(\mathrm{e}^{x_j}-\mathrm{e}^k)q(x_j)] \quad (22.8)$$

$$=\frac{\Delta x}{2}[\mathrm{e}^{x_{j+1}}-\mathrm{e}^k)q(x_{j+1})+(\mathrm{e}^{x_j}-\mathrm{e}^k)q(x_j)] \quad (22.9)$$

代回(22.7)，可得通过梯形公式得到的以下表达式

$$\widehat{C}(k) \approx \mathrm{e}^{-rT}\frac{\Delta x}{2}[1(\mathrm{e}^{x_0-\mathrm{e}^k})q(x_0)+2(\mathrm{e}^{x_1}-\mathrm{e}^k)q(x_1)+\cdots \quad (22.10)$$

$$+2(\mathrm{e}^{x_{n-1}}-\mathrm{e}^k)q(x_{n-1})+1(\mathrm{e}^{x_n}-\mathrm{e}^k)q(x_n)] \quad (22.11)$$

在辛普森公式中，我们要一次考虑两个相邻小区间，[⊖]要找到一个经过这 3 个端点的二次函数，然后用它来代替原函数，并求积分，即

$$\int_{x_j}^{x_{j+2}}(\mathrm{e}^s-\mathrm{e}^k)q(s)\mathrm{d}s \quad (22.12)$$

$$\approx\frac{x_{j+2}-x_j}{6}[(\mathrm{e}^{x_j}-\mathrm{e}^k)q(x_j)+4(\mathrm{e}^{x_{j+1}}-\mathrm{e}^k)q(x_{j+1})+(\mathrm{e}^{x_{j+2}}-\mathrm{e}^k)q(x_{j+2})] \quad (22.13)$$

$$=\frac{\Delta x}{3}[(\mathrm{e}^{x_j}-\mathrm{e}^k)q(x_j)+4(\mathrm{e}^{x_{j+1}}-\mathrm{e}^k)q(x_{j+1})+(\mathrm{e}^{x_{j+2}}-\mathrm{e}^k)q(x_{j+2})] \quad (22.14)$$

代回(22.7)，可得通过辛普森公式得到的以下表达式

$$\widehat{C}(k) \approx \mathrm{e}^{-rT}\frac{\Delta x}{3}[1(\mathrm{e}^{x_0}-\mathrm{e}^k)q(x_0)+4(\mathrm{e}^{x_1}-\mathrm{e}^k)q(x_1)$$

$$+2(\mathrm{e}^{x_2}-\mathrm{e}^k)q(x_2)+\cdots \quad (22.15)$$

$$+2(\mathrm{e}^{x_{n-2}}-\mathrm{e}^k)q(x_{n-2})+4(\mathrm{e}^{x_{n-1}}-\mathrm{e}^k)q(x_{n-1})+1(\mathrm{e}^{x_n}-\mathrm{e}^k)q(x_n)] \quad (22.16)$$

这两种方法的阶都是 $O(n)$. 注意，这种近似计算方法与金融模型本身之间是独立的 .

为了了解这两种方法是如何运作的，我们以一个欧式看涨期权定价模型作为例子 . 362

例 1 几何布朗运动/Black-Scholes 模型

假设某股票价格服从集合布朗运动：

$$\mathrm{d}S_t=(r-q)S_t\mathrm{d}t+\sigma S_t\mathrm{d}W_t \quad (22.17)$$

其中 r 为无风险利率，q 为连续红利率，σ 为瞬时波动率 . 根据我们前面介绍的内容，根据伊藤引理，该随机微分方程的解为

$$S_t=S_0\mathrm{e}^{(r-q-\frac{\sigma^2}{2})t+\sigma W_t} \quad (22.18)$$

$$=S_0\mathrm{e}^{(r-q-\frac{\sigma^2}{2})t+\sigma\sqrt{t}Z} \quad (22.19)$$

$$=\mathrm{e}^{\ln S_0+(r-q-\frac{\sigma^2}{2})t+\sigma\sqrt{t}Z} \quad (22.20)$$

⊖ 这意味着共有 $\frac{n}{2}$ 项，因此须将 n 设为偶数 .

其中 Z 为标准正态随机变量，意味着 S_t 服从对数正态分布，均值为 $\ln S_0+\left(r-q-\frac{\sigma^2}{2}\right)t$，标准差为 $\sigma\sqrt{t}$，即

$$f(S_T \mid S_0)=\frac{e^{-\frac{1}{2}\left(\frac{\ln S_T-\ln S_0-(r-q-\sigma^2/2)T}{\sigma\sqrt{T}}\right)^2}}{\sigma S_T\sqrt{2\pi T}} \tag{22.21}$$

因此看涨期权价格为

$$C(K)=\int_K^{\infty} e^{-rT}(S_T-K)f(S_T \mid S_0)dS_T \tag{22.22}$$

$$=\int_K^{\infty} e^{-rT}(S_T-K)\times\frac{e^{-\frac{1}{2}\left(\frac{\ln S_T-\ln S_0-(r-q-\sigma^2/2)T}{\sigma\sqrt{T}}\right)^2}}{\sigma S_T\sqrt{2\pi T}}dS_T \tag{22.23}$$

$$=\int_K^{\infty} e^{-rT}(e^{s_T}-e^k)\times\frac{e^{-\frac{1}{2}\left(\frac{\ln s_T-\ln s_0-(r-q-\sigma^2/2)T}{\sigma\sqrt{T}}\right)^2}}{\sigma\sqrt{2\pi T}}dS_T \tag{22.24}$$

$$=e^{-rT}\int_K^{\infty}(e^{s_T}-e^k)q(S_T \mid S_0)dS_T=\hat{C}(k) \tag{22.25}$$

在(22.24)中我们使用了变量代换 $s_T=\ln(S_T)$，$k=\ln(K)$. ◀

例 2 Black-Scholes 欧式期权数值定价方法

我们设定参数为：当前价格 $S_0=100$，执行价格 $K=100$，到期日 $T=1$，波动率 $\sigma=30\%$，无风险利率 $r=2\%$，红利率 $q=1.25\%$，该看涨期权的 Black-Scholes 价格为 12.1040. 如果使用刚才介绍的梯形公式和辛普森公式，我们得到的数值结果如表 22-1 所示. 和预期一样，数值结果对 $b=\ln B$ 的大小很敏感. 当 b 较大时，两种方法(尤其是第二种)的结果都很接近真实解，几乎与 n 无关.

该方法有一个主要问题. 对于许多随机过程，封闭形式的条件分布函数无法获得，或者很难计算. 不过也有许多随机过程，虽然无法求出分布函数，它们的特征函数的封闭形式已知. ◀

首先，我们给出随机过程特征函数的定义，并以几何布朗运动为例，介绍求特征函数的方法. 然后介绍如何根据特征函数对欧式看涨期权定价.

定义 24 傅里叶变换

定义函数 $f(x)$ 的傅里叶变换为：

$$\phi(\nu)=\int_{-\infty}^{\infty} e^{i\nu x}f(x)dx \tag{22.26}$$

定义 25 傅里叶逆变换

已知函数 $f(x)$ 的傅里叶变换 $\varphi(\nu)$，则我们可以通过傅里叶逆变换得到函数 $f(x)$：

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\nu x}\phi(\nu)d\nu \tag{22.27}$$

表 22-1 对不同值的 B 和 n 使用梯形或辛普森公式得 Black-Scholes 价格
(Black-Scholes 价格为 12.1040)

B	n	数值积分路径	
		梯形公式	辛普森公式
200	2^4	11.1868	11.0992
	2^5	11.2100	11.1637
	2^6	11.2154	11.1917
	2^7	11.2167	11.2047
	2^8	11.2170	11.2110
	2^{10}	11.2171	11.2156
400	2^4	12.0111	12.1053
	2^5	12.0821	12.1037
	2^6	12.0984	12.1037
	2^7	12.1024	12.1037
	2^8	12.1033	12.1037
	2^{10}	12.1037	12.1037
800	2^4	11.1868	12.1141
	2^5	12.0554	12.1045
	2^6	12.0922	12.1040
	2^7	12.1011	12.1040
	2^8	12.1033	12.1040
	2^{10}	12.1040	12.1040

定义 26 特征函数

设 $f(x)$是某随机变量 x 的概率密度函数，满足 $f(x)\geqslant 0$，$\int f(x)=1$. 其傅里叶变换的结果叫做特征函数：

$$\phi(\nu)=\int_{-\infty}^{\infty} \mathrm{e}^{i\nu x} f(x)\mathrm{d}x \tag{22.28}$$

$$= E(\mathrm{e}^{i\nu x}) \tag{22.29}$$

很明显，通过对特征函数作傅里叶逆变换，即可得到密度函数. 注意，特征函数在 $-iv$ 的取值恰好就是矩母函数. 因此，我们可以通过特征函数求随机变量的各阶矩.

例 3 标准正态分布的特征函数

金融数学中最常用的分布是标准正态分布. 它是扩散过程的主要成分，也是我们这里讨论的大多数模型的核心. 如果 $Z\sim\mathcal{N}(0, 1)$，那么它的特征函数为

$$\phi_Z(\nu) = E(\mathrm{e}^{i\nu Z}) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\exp\left(i\nu z - \frac{1}{2}z^2\right)\mathrm{d}z \tag{22.30}$$

我们首先考虑以下积分

$$E(\mathrm{e}^{sZ}) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\exp\left(sz - \frac{1}{2}z^2\right)\mathrm{d}z \tag{22.31}$$

将括号内的项凑成完全平方，可得

$$\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left(sz-\frac{1}{2}z^2\right)dz=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}(z^2-2sz)\right)dz \tag{22.32}$$

$$=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}(z-s)^2+\frac{1}{2}s^2\right)dz \tag{22.33}$$

$$=\frac{1}{\sqrt{2\pi}}e^{\frac{1}{2}s^2}\int_{-\infty}^{\infty}\exp\left(-\frac{1}{2}(z-s)^2\right)dz \tag{22.34}$$

利用

$$\int_{-\infty}^{\infty}e^{-\frac{1}{2}u^2du}=\sqrt{2\pi} \tag{22.35}$$

可得

$$E(e^{sZ})=\exp\left(\frac{s^2}{2}\right) \tag{22.36}$$

根据 Grimmett 和 Stirzaker(1992)中的内容，用 $i\nu$ 替代 s 可得

364

$$\phi_Z(\nu)=E(e^{i\nu z})=e^{-\frac{\nu^2}{2}} \tag{22.37}$$

根据复变函数的解析连续理论，这是可行的． ◀

例 4 正态分布的特征函数

均值为 μ、方差为 σ 的正态随机变量可以通过对标准正态分布处理后获得，即利用 $X=\mu+\sigma Z$，使得 $X\sim\mathcal{N}(\mu,\sigma^2)$．所以，它的特征函数为

$$\phi_X(\nu)=E(e^{i\nu X}) \tag{22.38}$$

$$=E(e^{i\nu(\mu+\sigma Z)}) \tag{22.39}$$

$$=e^{i\nu\mu}E(e^{i\nu\sigma Z}) \tag{22.40}$$

$$=e^{i\nu\mu}e^{-\frac{(\sigma\nu)^2}{2}} \tag{22.41}$$

$$=e^{i\mu\nu-\frac{\sigma^2\nu^2}{2}} \tag{22.42}$$

布朗运动 W_t 是许多资产定价模型的关键成分．我们知道

$$W_t-W_0=W_t\sim\mathcal{N}(0,t) \tag{22.43}$$

因此如果 $X_t=W_t$，它的特征函数为

$$E(e^{i\nu X_t})=E(e^{i\nu W_t})=E(e^{i\nu\sqrt{t}Z})=e^{-\frac{\nu^2t}{2}} \tag{22.44}$$ ◀

例 5 几何布朗运动的对数的特征函数

金融数学中最常用的资产定价模型是几何布朗运动．该模型现在仍是数量金融的关键工具，并且是许多现代衍生品定价模型的衡量标准．它的发明极大地促进了期权市场的壮大以及流动性的增强．本例将简要介绍该模型的由来以及它的随机微分方程、特征函数．

当我们对资产价格建模时，假定标的过程为资产价格 S_t，在 t 时刻，将满足以下随机微分方程，即我们熟知的 Black-Scholes 随机微分方程：

$$dS_t=(r-q)S_tdt+\sigma S_tdW_t \tag{22.45}$$

根据该方程，资产价格对数的增长率为 $r-q$，波动率为 σ.

根据伊藤引理，该随机微分方程的解为

$$S_T = S_0 e^{\left(r-q-\frac{\sigma^2}{2}\right)T+\sigma W_T} \tag{22.46}$$

$$= S_0 e^{\left(r-q-\frac{\sigma^2}{2}\right)T+\sigma\sqrt{T}Z} \tag{22.47}$$

或者等价地

$$s_T = \ln S_T = s_0 + \left(r-q-\frac{\sigma^2}{2}\right)T + \sigma\sqrt{T}Z \tag{22.48}$$

其中 $s_0 = \ln S_0$，因此

$$s_T \sim \mathcal{N}\left(s_0 + \left(r-q-\frac{\sigma^2}{2}\right)T, \sigma^2 T\right) \tag{22.49}$$

利用正态随机变量的特征函数，我们很容易就能求出资产价格对数的特征函数：

$$\phi_{s_T}(\nu) = e^{i\left(s_0+\left(r-q-\frac{\sigma^2}{2}\right)T\right)\nu-\frac{\sigma^2\nu^2}{2}T} \tag{22.50}$$

22.1 用傅里叶变换进行衍生品定价

本章我们将讨论傅里叶变换在衍生品定价中的使用．事实上，资产价格分布的特征函数就是其概率分布函数的傅里叶变换．因此分布函数可以通过对特征函数进行傅里叶逆变换得到．对于很多只知道特征函数封闭形式表达式的模型来说，这一点尤为重要．我们将简要介绍衍生品定价中的各种变换方法，着重介绍快速傅里叶变换(FFT)的相关内容．在衍生品定价中使用傅里叶变换的第一个重大进展来自 Carr 和 Madan(1999)．虽然这个技巧是一项可观的突破，像很多其他方法一样，快速傅里叶变换也有一些需要权衡的地方．当我们的模型已知特征函数，尤其是有些时候仅能得到特征函数的时候，傅里叶变换非常有用，能有效地帮助我们定价．而且，傅里叶逆变换的速度非常快，能以 $O(n\ln(n))$ 的速度解决衍生品定价问题．此外，该方法不但能帮我们以 $O(n\ln(n))$ 的速度快速计算期权价格，还能计算 n 种不同执行价格的期权．由于对期权价格的计算还存在一些限制条件，我们可以从该方法中提取比一般方法更多的信息，这在计算中很重要．

365

并不是所有衍生品都能用本方法定价．特别地，该方法只限于支付为欧式的金融衍生品，即支付与路径完全无关的衍生品．此外，该方法的推导十分依赖于支付类型：在 Carr 和 Madan 的原始文献中，作者只介绍了两种特殊的支付类型．因此我们会被限制在一个较小但很重要的衍生品支付的子集里．为了使该方法有效，我们需要定义一个阻尼因子 α，其最优值必须确定下来．最后，如果期权虚值程度很高，那该方法的精确程度会下降．

22.1.1 用傅里叶变换对看涨期权定价

设 $\phi(\nu)$ 为标的证券 s_T 的特征函数，即

$$\phi(\nu) = \int_{-\infty}^{\infty} e^{i\nu s_T} q_{s_T} \, ds_T \tag{22.51}$$

欧式看涨期权价格 $C(K)$ 可以表示为：

$$C(K)=\mathrm{e}^{-rT}E_t[(S_T-K)^+] \tag{22.52}$$

$$=\mathrm{e}^{-rT}\int_k^{\infty}(S_T-K)f(S_T \mid S_0)\,\mathrm{d}S_T \tag{22.53}$$

$$=\mathrm{e}^{-rT}\int_k^{\infty}(\mathrm{e}^{s_T}-\mathrm{e}^k)q(s_T)\,\mathrm{d}s_T \tag{22.54}$$

$$=\mathrm{e}^{-rT}\int_k^{\infty}(\mathrm{e}^{s}-\mathrm{e}^k)q(s)\,\mathrm{d}s \tag{22.55}$$

$$=\widehat{C}(k) \tag{22.56}$$

为了简化起见，我们在最后一个积分方程省略了下标 T. 既然我们已经有了期权价格$\widehat{C}(k)$关于对数价格密度的表达式，我们能够通过它来计算傅里叶变换，从而与特征函数建立联系. 我们将傅里叶变换的结果定义为 $\psi(\nu)$，即

$$\psi(\nu)=\int_{-\infty}^{\infty}\mathrm{e}^{i\nu k}\widehat{C}(k)\,\mathrm{d}k \tag{22.57}$$

$$=\int_{-\infty}^{\infty}\mathrm{e}^{i\nu k}\left(\mathrm{e}^{-rT}\int_k^{\infty}(\mathrm{e}^{s}-\mathrm{e}^k)q(s)\,\mathrm{d}s\right)\mathrm{d}k \tag{22.58}$$

$$=\mathrm{e}^{-rT}\int_{-\infty}^{\infty}\mathrm{e}^{i\nu k}\left(\int_k^{\infty}(\mathrm{e}^{s}-\mathrm{e}^k)q(s)\,\mathrm{d}s\right)\mathrm{d}k \tag{22.59}$$

为了求积分，我们需要改变积分顺序. 此时我们需要使用 Fubini 定理，从而改变重积分的积分顺序.

$$\psi(\nu)=\mathrm{e}^{-rT}\int_{-\infty}^{\infty}\int_{-\infty}^{s}\mathrm{e}^{i\nu k}(\mathrm{e}^{s}-\mathrm{e}^k)q(s)\,\mathrm{d}k\,\mathrm{d}s \tag{22.60}$$

$$=\mathrm{e}^{-rT}\int_{-\infty}^{\infty}q(s)\left(\int_{-\infty}^{s}\mathrm{e}^{i\nu k}(\mathrm{e}^{s}-\mathrm{e}^k)\,\mathrm{d}k\right)\mathrm{d}s \tag{22.61}$$

现在我们可以计算内层的那个积分

$$\int_{-\infty}^{s}\mathrm{e}^{i\nu k}(\mathrm{e}^{s}-\mathrm{e}^k)\,\mathrm{d}k=\int_{-\infty}^{s}\mathrm{e}^{i\nu k}\mathrm{e}^{s}\,\mathrm{d}k-\int_{-\infty}^{s}\mathrm{e}^{i\nu k}\mathrm{e}^{k}\,\mathrm{d}k=\mathrm{e}^{s}\left.\frac{\mathrm{e}^{i\nu k}}{i\nu}\right|_{-\infty}^{s}-\left.\frac{\mathrm{e}^{(i\nu+1)k}}{i\nu+1}\right|_{-\infty}^{s} \tag{22.62}$$

366

我们可以看到第一个积分并不收敛，第一项是没有定义的. 为了确保这个方法有效，我们需要对问题重新进行公式化，根据 Carr 和 Madan(1999)，我们定义

$$c(k)=\mathrm{e}^{\alpha k}\widehat{C}(k) \tag{22.63}$$

即对期权价格乘以执行价格的指数函数. 该项会成为内层积分的阻尼成分，使原来的项收敛，使得傅里叶变换可以计算. 我们重新定义修正期权价格 $c(k)$的特征函数 $\psi(\nu)$，其推导则变成了：

$$\psi(\nu)=\int_{-\infty}^{\infty}\mathrm{e}^{i\nu k}c(k)\,\mathrm{d}k$$

$$=\int_{-\infty}^{\infty}\mathrm{e}^{i\nu k}\left(\mathrm{e}^{-rT}\mathrm{e}^{\alpha k}\int_k^{\infty}(\mathrm{e}^{s}-\mathrm{e}^k)q(s)\,\mathrm{d}s\right)\mathrm{d}k$$

$$= e^{-rT}\int_{-\infty}^{\infty}\int_{-\infty}^{s} e^{(\alpha+i\nu)k}(e^{s}-e^{k})q(s)\mathrm{d}k\mathrm{d}s$$

$$= e^{-rT}\int_{-\infty}^{\infty} q(s)\left(\int_{-\infty}^{s} e^{(\alpha+i\nu)k}(e^{s}-e^{k})\mathrm{d}k\right)\mathrm{d}s \tag{22.64}$$

有了阻尼因子，内层的积分变得收敛

$$\int_{-\infty}^{s} e^{\alpha k}e^{i\nu k}(e^{s}-e^{k})\mathrm{d}k = \int_{-\infty}^{s} e^{(\alpha+i\nu)k}(e^{s}-e^{k})\mathrm{d}k \tag{22.65}$$

$$= e^{s}\left.\frac{e^{(\alpha+i\nu)k}}{(\alpha+i\nu)}\right|_{-\infty}^{s} - \left.\frac{e^{(\alpha+i\nu+1)k}}{(\alpha+i\nu+1)}\right|_{-\infty}^{s} \tag{22.66}$$

由于 α 大于 0，在负无穷处这两项都会消失，因此我们有

$$\int_{-\infty}^{s} e^{(\alpha+i\nu)k}(e^{s}-e^{k})\mathrm{d}k = e^{s}\frac{e^{(\alpha+i\nu)s}}{\alpha+i\nu}-\frac{e^{(\alpha+i\nu+1)s}}{\alpha+i\nu+1} = \frac{e^{(\alpha+i\nu+1)s}}{(\alpha+i\nu)(\alpha+i\nu+1)} \tag{22.67}$$

将(22.67)带入(22.64)，即得修正期权价格的特征函数表达式

$$\psi(\nu) = e^{-rT}\int_{-\infty}^{\infty} q(s)\frac{e^{(\alpha+i\nu+1)s}}{(\alpha+i\nu)(\alpha+i\nu+1)}\mathrm{d}s \tag{22.68}$$

将于 s 无关的项提取出来，可得

$$\psi(\nu) = \frac{e^{-rT}}{(\alpha+i\nu)(\alpha+i\nu+1)}\int_{-\infty}^{\infty} e^{(\alpha+i\nu+1)s}q(s)\mathrm{d}s \tag{22.69}$$

将积分中的 i 提到括号外面，我们可以发现该积分正是资产价格对数的特征函数在 $(\nu-(\alpha+1)i)$ 处的值，即

$$\psi(\nu) = \frac{e^{-rT}}{(\alpha+i\nu)(\alpha+i\nu+1)}\int_{-\infty}^{\infty} e^{i(\nu-(\alpha+1)i)s}q(s)\mathrm{d}s \tag{22.70}$$

$$= \frac{e^{-rT}}{(\alpha+i\nu)(\alpha+i\nu+1)}\phi(\nu-(\alpha+1)i) \tag{22.71}$$

如果知道标的证券价格对数的特征函数 $\phi(\nu)$，我们就可以计算 $\psi(\nu)$. 修正看涨期权价格的傅里叶变换为

$$\psi(\nu) = \int_{-\infty}^{\infty} e^{i\nu k}c_{T}(k)\mathrm{d}k \tag{22.72}$$

$$= \frac{e^{-rT}}{(\alpha+i\nu)(\alpha+i\nu+1)}\phi(\nu-(\alpha+1)i) \tag{22.73}$$

由于我们已经知道修正看涨期权价格的特征函数

$$\psi(\nu) = \int_{-\infty}^{\infty} e^{i\nu k}c_{T}(k)\mathrm{d}k \tag{22.74}$$

$$= \int_{-\infty}^{\infty} e^{i\nu k}e^{\alpha k}C_{T}(k)\mathrm{d}k \tag{22.75}$$

367

我们可以利用傅里叶逆变换得到

$$\hat{C}(k) = \frac{e^{-\alpha k}}{2\pi}\int_{-\infty}^{\infty} e^{-i\nu k}\psi(\nu)\mathrm{d}\nu \tag{22.76}$$

然而，$\hat{C}(k)$应该是一个实数．$\psi(\nu)$的偶数项为实数，而奇数项为虚数．所以我们只需关心

期权价格的实数部分．因此我们把它当作一个偶函数，可得

$$\hat{C}(k)=\frac{e^{-\alpha k}}{\pi}\int_0^{\infty} e^{-i\nu k}\psi(\nu)d\nu \tag{22.77}$$

即看涨期权价格．利用方程(22.73)以及 $\psi(\nu)$的傅里叶逆变换，我们可以计算 $\hat{C}(k)$.

$$\hat{C}(k)=\frac{e^{-\alpha k}}{\pi}\int_0^{\infty} e^{-i\nu k}\psi(\nu)d\nu \tag{22.78}$$

其中 $\psi(\nu)$是将要确定的已知函数，α 是某个大于零的合适的参数．

22.1.2 计算定价积分

截至目前的介绍，傅里叶变换能帮助我们在特征函数已知而分布函数未知的情况下计算期权价格．但是我们最后需要求解一个积分．如果我们不能求出该积分，那么该方法的必要性也就无从体现了．注意

$$\hat{C}(k)=\frac{e^{-\alpha k}}{\pi}\int_0^{\infty} e^{-i\nu k}\psi(\nu)d\nu \tag{22.79}$$

该积分可以通过数值积分方法很容易地计算．首先，我们对它进行估计，定义 B 作为积分范围的上界．我们可以用梯形公式进行估算．设 n 为小区间的数量，相邻端点的距离为 $\Delta\nu=\frac{b}{n}=\eta$，$\nu_j=(j-1)\eta$为末端点的位置．运用梯形公式，可得

$$\hat{C}(k)=\frac{e^{-\alpha k}}{\pi}\int_0^{\infty} e^{-i\nu k}\psi(\nu)d\nu \tag{22.80}$$

$$\approx\frac{e^{-\alpha k}}{\pi}\int_0^{\infty} e^{-i\nu k}\psi(\nu)d\nu \tag{22.81}$$

$$\approx\frac{e^{-\alpha k}}{\pi}(e^{-i\nu_1 k}\psi(\nu_1)+2e^{-i\nu_2 k}\psi(\nu_2)+\cdots+2e^{-i\nu_N k}\psi(\nu_N) \tag{22.82}$$

$$+e^{-i\nu_{N+1}k}\psi(\nu_{N+1}))\frac{\eta}{2} \tag{22.83}$$

由于各项以指数级减小，我们可以省略最后一项，近似为

$$C(k)\approx\frac{e^{-\alpha k}}{\pi}\sum_{j=1}^{n} e^{-i\nu_j k}\psi(\nu_j)\omega_j \tag{22.84}$$

其中 $\omega_j=\frac{\eta}{2}(2-\delta_{j-1})$. 如果需要更精确的结果，我们可以使用辛普森公式，最终将得到：

$$\hat{C}(k)\approx\frac{e^{-\alpha k}}{\pi}\sum_{j=1}^{n} e^{-i\nu_j k}\psi(\nu_j)\omega_j \tag{22.85}$$

其中 $\omega_j=\frac{\eta}{3}(3+(-1)^j-\delta_{j-n})$. 这里 $\delta_0=1$，$\delta_j=0$，$j\geqslant 1$.

快速傅里叶变换

如果可以直接积分，那么估计定价积分就不是一种很好的方法．Cooley 和 Tukey (1965)提出的快速傅里叶变换(FFT)算法以及其他人的进一步研究成果对于一类离散傅里

368

叶逆变换提供了一类有效算法，其中样本点是 2 的指数次．这类变换有以下形式

$$\omega(m)=\sum_{j=1}^{n}\mathrm{e}^{-i\frac{2\pi}{n}(j-1)(m-1)}x(j),\ m=1,\cdots,n \tag{22.86}$$

这些方程每次逆变换作 n 次乘法，总共作 n^2 次乘法，而 Cooley-Tukey FFT 算法将计算量缩减到了 $n\log(n)$次乘法，他们使用了分治算法，从而分解离散傅里叶变换．这对于估计傅里叶积分非常关键，因为它能够大大加快我们在 FFT 方法下对期权价格的计算．

为了将我们的期权定价公式转化为 FFT 形式，我们首先在执行价格附近建立一个执行价格的范围，涵盖所有我们想要计算的情况．典型的例子是关于某特定标的物的平价期权，我们定义执行价格对数的范围为

$$k_m=\beta+(m-1)\Delta k=\beta+(m-1)\lambda,\quad m=1,\cdots,n \tag{22.87}$$

其中 $\beta=\ln S_0-\frac{\lambda n}{2}$，它确保该平价期权的执行价格正好在整个范围的中间．对于 $\hat{C}(m)$，现在我们有

$$\hat{C}(k_m)\approx\frac{\mathrm{e}^{-\alpha k_m}}{\pi}\sum_{j=1}^{n}\mathrm{e}^{-i\nu_j k_m}\psi(\nu_j)\omega_j,\quad m=1,\cdots,n \tag{22.88}$$

$$=\frac{\mathrm{e}^{-\alpha k_m}}{\pi}\sum_{j=1}^{n}\mathrm{e}^{-i\nu_j(\beta+(m-1)\Delta k)}\psi(\nu_j)\omega_j \tag{22.89}$$

$$=\frac{\mathrm{e}^{-\alpha k_m}}{\pi}\sum_{j=1}^{n}\mathrm{e}^{-i(j-1)\eta(m-1)\lambda}\mathrm{e}^{-i\beta\nu_j}\psi(\nu_j)\omega_j \tag{22.90}$$

$$=\frac{\mathrm{e}^{-\alpha k_m}}{\pi}\sum_{j=1}^{n}\mathrm{e}^{-i\lambda\eta(j-1)(m-1)}\mathrm{e}^{-i\beta\nu_j}\psi(\nu_j)\omega_j \tag{22.91}$$

如果令

$$\lambda\eta=\frac{2\pi}{n} \tag{22.92}$$

以及

$$x(j)=\mathrm{e}^{-i\beta\nu_j}\psi(\nu_j)\omega_j \tag{22.93}$$

我们可以得到(22.86)中的 FFT 原始形式．现在我们在 FFT 中仅需使用 $O(n\log n)$次乘法，就可以产生 n 个期权价格．虽然相比于用直接积分的 $O(n)$次乘法计算一个单独的期权价格，整个计算过程速度更慢，但实际上关于一个标的物，我们很少只计算一个单独的期权. FFT 方法具有明显的优势．$O(n\log n)$次乘法平均到 n 个期权上，每个期权只需要 $O(\log n)$次乘法．

但是，这 n 个期权很可能不会是生成隐含波动率平面所需要的 n 个期权．因为 FFT 方法中的执行价格是由

$$k_m=\beta+(m-1)\Delta k=\beta+(m-1)\lambda \tag{22.94}$$

所定义的．我们可以通过修正 n 来修正 $\eta=\frac{b}{n}$，从而改变用 FFT 求期权价格时所使用的执

行价格. 这使得我们可以获取足够的信息来生成波动率平面，同时又比直接积分花费更少的时间，后者计算 n 个期权价格需要 $O(n^2)$ 次乘法.

22.1.3 快速傅里叶变换的使用

假设我们已知标的过程 X_t 的对数的特征函数 $\phi(\nu)$. 这里我们取 η 和 $n=2^l$，计算出 $\lambda=\frac{2\pi}{n\eta}$，$\nu_j=(j-1)\eta$并设定 α. 构造向量 $\boldsymbol{x}$

$$\boldsymbol{x}=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=\begin{pmatrix}\frac{\eta}{2}\frac{e^{-rT}}{(\alpha+i\nu_1)(\alpha+i\nu_1+1)}e^{-i(\ln S_0-\frac{\lambda n}{2})\nu_1}\phi(\nu_1-(\alpha+1)i)\\ n\frac{e^{-rT}}{(\alpha+i\nu_2)(\alpha+i\nu_2+1)}e^{-i(\ln S_0-\frac{\lambda n}{2})\nu_2}\phi(\nu_2-(\alpha+1)i)\\ \vdots\\ n\frac{e^{-rT}}{(\alpha+i\nu_n)(\alpha+i\nu_n+1)}e^{-i(\ln S_0-\frac{\lambda n}{2})\nu_n}\phi(\nu_n-(\alpha+1)i)\end{pmatrix} \tag{22.95}$$

369

向量 $\boldsymbol{x}$ 是 FFT 的输入，而输出是结构与之相同的向量 $\boldsymbol{y}$(例如，在 MATLAB 中我们若想完成此步骤，只需要输入 $\boldsymbol{y}=\text{fft}(\boldsymbol{x})$)，执行价格为 $k_m(m=1,\cdots,n)$的看涨期权价格为

$$\begin{pmatrix}\hat{C}(k_1)\\\hat{C}(k_2)\\\vdots\\\hat{C}(k_n)\end{pmatrix}=\begin{pmatrix}\frac{e^{-\alpha(\ln S_0-\frac{n}{2}\lambda)}}{\pi}\text{Re}(y_1)\\ \frac{e^{-\alpha(\ln S_0-(\frac{n}{2}-1)\lambda)}}{\pi}\text{Re}(y_2)\\ \vdots\\ \frac{e^{-\alpha(\ln S_0-(\frac{n}{2}-(n-1))\lambda)}}{\pi}\text{Re}(y_n)\end{pmatrix} \tag{22.96}$$

其中 $\text{Re}(y_j)$是 y_j 的实部.

22.2 观察与发现

阻尼因子的使用使我们得以通过傅里叶变换对普通期权定价. 从表面上看，α 好像并没有参与积分的计算，而是藏在 $\psi(\nu)$中. 我们已经知道对于看涨期权 α 必须大于 0，而且似乎只要 α 大于 0 我们就能得到大致相同的结果. 然而，实际并不是这样. 我们可以证明结果对 α 的取值很敏感，并且可以找到 α 合适的取值范围，同时阐明它与随机模型的选择的关系.

我们可以发现 λ 与 η 具有以下关系

$$\lambda\eta=\frac{2\pi}{n} \tag{22.97}$$

且

$$\eta=\frac{b}{n} \tag{22.98}$$

这意味着涉及 4 个参数，n、b、λ 和 η. 只有两个可以独立取值，η 由 b 决定，而最后一个由(22.97)的限制条件决定．如果我们限制了固定的计算量，我们可以事先确定积分项的数量 n. 考虑以上假定，我们只有两个自由变量，即积分上界 b，和分割 $\log(K)$ 的子区间宽度$\lambda=\Delta k$，而它们是成反比的．因此我们就有了一个上界和执行价格步长之间的需要权衡的地方．前者决定了积分的精确度，而后者决定了我们计算的内容与市场信息是否对应．选择的 b 越大，我们的积分结果越精确；然而，如果我们想让样本点的间距固定为某一数值，以此来保证积分的准确性，我们必须要对 λ 施加限制，从而确定执行价格对数的间距．

22.3 小结

如果已知随机过程的概率密度函数，我们可以利用数值积分的方法，直接得到期权价格．

在大多数情况下，我们不知道密度函数的解析式或积分形式．但是，我们通常可以找到标的证券价格的特征函数，或者证券价格对数的特征函数的解析形式或半解析形式．根据 Carr 和 Madan(1999)如果我们已知特征函数的解析式，我们可以通过傅里叶变换有效地得到期权价格．按照其内容，我们先把欧式看涨期权的定价问题写成资产价格对数的密度函数形式，然后我们就可以用傅里叶变换求期权价格．

370

许多衍生工具，包括普通期权、利率的上/下限协议、互换期权，可以表示为简单的看涨或看跌期权的组合．因此，这种方法也可以用来对这些工具定价．

FFT 方法可以用于任何资产价格特征函数已知的模型．它适用于路径独立的欧式期权且最终支付限制于某个集合的情况．该方法速度较快，计算 n 个期权价格只需 $O(\log(n))$次乘法，并且可以一次性产生 n 个期权价格．存在的问题是，如果衍生品的支付结构发生了任何变化，整个推导都要重新进行．同时，它需要估测合适的 α，而当期权虚值程度很高时，它的结果将变得很不准确．

22.4 习题

1. 证明在用 FFT 进行欧式看跌期权定价时，我们需要令 $\alpha<-1$.

2. 在使用 FFT 时，我们定义执行价格对数的范围为 $k_m=\beta+(m-1)\Delta k=\beta+(m-1)\lambda(m=1,\cdots,N)$，$\beta$ 为某个合适的值．β 有很多选择．其中一种是令 $\beta=\ln S_0-\frac{\lambda N}{2}$. 这样选择 β 可以确保平价期权的执行价格恰好落在正中央，因为 S_0 是即期价格也是执行价格．我们感兴趣的是，如果改变 $k=\log K$，我们应该怎样选择 β?

(a) 如果我们想求某个具体执行价格 $k_0=\log K_0$ 时的期权佣金，我们该如何选择 β?

(b) 如何选择 β，才能使第一个元素与 k_0 对应的期权佣金一致?

3. 根据之前介绍，在 Black-Scholes 框架下，股票价格对数的特征函数为：

$$\phi(u)=E(e^{iu\ln S_T}) \tag{22.99}$$

$$=E(e^{ius_T}) \tag{22.100}$$

$$= \exp\left(i(s_0 + (r - q - \sigma^2/2)T)u - \frac{1}{2}\sigma^2 u^2 T\right) \tag{22.101}$$

使用 FFT 方法对欧式看涨期权定价，参数如下：即期价格 $S_0=100$，执行价格 $K=100$，无风险利率 $r=1.25\%$，红利率 $q=1.0\%$，期限 $T=2$ 年，波动率 $\sigma=20\%$. 并将结果与封闭形式的看涨期权价格比较．对 B、N、α 取不同的值．将结果与解析解比较，并分析对 B、N、α 的敏感程度．

4. 在 Heston 随机波动率模型中，股票价格服从以下随机微分方程：

$$\begin{aligned} dS_t &= rS_t dt + \sqrt{v_t} S_t dW_t^{(1)} \\ dv_t &= \kappa(\theta - v_t)dt + \sigma\sqrt{v_t} S_t dW_t^{(2)} \end{aligned} \tag{22.102}$$

其中两个布朗成分 $W_t^{(1)}$ 和 $W_t^{(2)}$ 可能相关且相关系数为 ρ. 变量 v_t 代表了均值回归的随机波动率，其中 θ 是长期方差，κ 是均值回归速度，σ 是方差的波动率．扩散项中 $\sqrt{v_t}$ 的存在使得波动率不会小于 0，同时波动率趋于 0 时扩散项也趋于 0. 根据 Hirsa(2012)，在 Heston 模型中，股票价格过程对数的特征函数为

$$\phi(u) = E(e^{iu\ln S_t}) \tag{22.103}$$

$$= \frac{\exp\left\{\dfrac{\kappa\theta t(\kappa - i\rho\sigma u)}{\sigma^2} + iutr + iu\ln S_0\right\}}{\left(\cosh\dfrac{\gamma t}{2} + \dfrac{\kappa - i\rho\sigma u}{\gamma}\sinh\dfrac{\gamma t}{2}\right)^{\frac{2\kappa\theta}{\sigma^2}}} \times \exp\left\{\frac{-(u^2+iu)v_0}{\gamma\coth\dfrac{\gamma t}{2} + \kappa - i\rho\sigma u}\right\} \tag{22.104}$$

371～372

其中 $\gamma = \sqrt{\sigma^2(\mu^2+iu) + (\kappa - i\rho\sigma u)^2}$. 利用 FFT 方法对欧式看跌期权定价，相关参数如下：即期价格 $S_0=100$，执行价格 $K=100$，期限 $T=2$ 年，无风险利率 $r=1.25\%$，波动率 $\sigma=20\%$，$\kappa=1$，$\theta=0.025$，$\rho=-0.7$，$\nu_0=0.05$. 对 B，N，α 取不同的值．

第 23 章　信用溢价和信用衍生品

本章涵盖了信用溢价和标准信用合约，总结了某些信用衍生品合约的建模和定价，包括违约相关性、可违约债券、信用违约互换以及一篮子违约互换．

23.1　标准合约

23.1.1　信用违约互换

信用违约互换(CDS)又称信用违约掉期，是最基础的信用衍生工具．信用违约互换是一种用于防范信用风险的保险．它是私下协商的双边合约．保护的购买者向出售者支付预付款，并在一段时间内支付固定的费用(或保费)．如果发生了某一个属于约定范畴的信用事件，保护的出售方须向购买方支付补偿．图 23-1 描述了一个 CDS 合约的支付结构：在该表中，X 方是违约保护的购买方，这意味着他拥有参考信用的空头．Y 方是违约保护的出售方，这意味着他拥有参考信用的多头．在合约的开始，X 向 Y 支付一笔预付款，并且在参考信用延续期间定期支付费用．如果发生违约，Y 向 X 支付 $1-R$. R 是债券的回收率．回收率表示债券在刚刚违约后的价格与面值的比率．它是一个介于 0 和 1 的数．而该债券是不受保证的现金流．



图 23-1　CDS

触发支付的信用事件可能是公司(参考实体)的破产，也可能是债券的违约，或者参考实体的债务问题．如果在互换期间没有发生信用事件，保护的购买方将持续支付保费直到到期日．而如果在到期之前发生了信用事件，保护的卖出方就需要向买入方支付补偿，从而补偿了保护买入方的经济损失．通常在合约开始时双方没有资金的交换，但他们在合约期间进行支付，因此这种行为叫做信用违约互换．

CDS 的另一个应用是一种叫做合成债务抵押债券(synthetic CDO)的证券化产品，我们会在后面介绍．

1. 当前全球 CDS 市场的状况如何？

CDS 市场起源于银行与其客户的私下洽商的协议．可能由于此类产品是在场外进行交易，对于它的具体诞生时间并没有明确记载．

根据美国财政部金融局(OCC)的记录，截至 2013 年第二季度，利率和货币衍生品的总面值为 217.12 万亿美元，而信用违约互换的总面值为 11.824 万亿美元(截至 2013 年 9 月 27 日)，约占总衍生品市场的 5%．截至 2013 年第二季度，全球范围内信用衍生品(13.382 万亿)占总衍生品市场的 5.72%．从增长的角度看，现在信用衍生品市场规模已经超过了实体衍生品市场，后者截至 2013 年第二季度为 2.08 万亿(数据来自 OCC).

CDS 最大的买入方是商业银行．在传统上，银行的业务是进行放贷，同时承担借款人的违约风险．CDS 市场为银行提供了很诱人的服务，使银行既可以规避风险，又不用从资产负债表上转移资产或者影响到借款人．此外，银行可以利用 CDS 使得原本只是集中于某374些产业或地域的投资组合更加多样化．银行是信用衍生品的净购买者，而保险公司逐渐成为 CDS 市场的主要参与者，它们出售信用保护，保障买方的投资收益．

截至 2012 年底该年 CDO 总价值为 1795 亿美元．截至 2013 年 10 月 1 日该年 CDO 总价值为 1507.13 亿美元(数据来自 SIFMA)．参与者主要是保险公司、金融担保人、对冲基金．

CDS 合约的到期日和规模不存在限制．但是，大多数合约名义金额在 1000 万到 2000 万美元之间，期限通常在 1 到 10 年，以 5 年期最为常见．单一 CDS 的名义金额的平均数和中位数分别是 1800 万美元和 1000 万美元(数据来自 NY Fed)．

2. 为什么信用违约互换能够吸引投资者？

信用违约互换可以用来对冲或者管理所面临的信用风险．它具有以下几个重要特征：

CDS 和参考实体的企业债券比较类似，但存在几个重要区别．CDS 不需要初始资金，可以作为杠杆头寸．当信用实体某个期限的现金债券还不存在时，就可以建立 CDS 协议．此外，市场参与者可以通过以买方身份进入 CDS 协议来建立参考信用的空头仓位．由于这些特点，CDS 是多样化投资组合和对冲风险的优秀工具．

近年来，为得到单一信用风险保护而进入 CDS 交易的投资者，越来越多地基于 CDS 指数进行交易．道琼斯 CDX 指数是一系列受投资者关注的指数．CDS 指数由 100 到 125 个流动性最高的 CDS 决定，权重平均，敞口多样化．

3. 到底什么是信用事件？

在 CDS 市场发展的早期，由于缺乏广泛接受的标准化文件，市场存在许多问题，而这是由于合约的期限和条款不够具体，存在许多盲点和技术漏洞．当信用事件发生，CDS 的买卖双方经常对合约的具体条款和期限产生争议．问题在于信用保护的买方希望保护的范围越大越好，而卖方恰恰相反．这很容易理解，因为 CDS 就像保险合同，保险的买方希望尽可能地在保险范围上多争取一点以获得索赔，而保险公司则希望尽量拒绝索赔或者尽量少赔．标准化文件的缺失严重阻碍了 CDS 市场的发展．直到 1999 年，ISDA 为信用衍生品合约颁布了一套新的主协定，其中包括了一系列关于信用衍生品的规定，填补了其在标准化文件上的缺失．后来，ISDA 颁布了 2003 ISDA 信用衍生品定义，以及 2002 主协定，进一步为出现的问题提供准则．

市场参与者普遍将以下 3 点视为信用事件最重要的标志：

- 破产——最明显的事件，意味着信用实体倒闭，无法偿还债务．
- 未能付款——当信用实体在一段宽限期后仍无法支付本金或利息．
- 375债务重组——这意味着债务结构及条款发生对债权人不利的变化．

债务重组可能是信用事件的一种，并且可能带来许多问题．重组使得可交割债券的剩余期限存在一定限制，以此能够控制最便宜可交割债券的价值．根据当前的 ISDA 协定，关于重组共有 4 种选择：

无重组——这种设定直接从合约中排除了重组．在允许重组的情况下，有时候保护的卖方可能会遇到一些轻度的信用事件，此时保护的买方并不一定会受到损失．而在当前不允许重组的设定下，这种情况不会发生．

完全重组——这种设定允许保护的购买方在任何形式的债务重组后交割任何期限的债券．

修正重组——修正重组将交付义务限定在重组后30个月之内到期的债券．这在美国和加拿大比较流行．

改良的修正重组——这是修正重组的改良版．它的产生是由于人们认为修正重组对交付义务的要求过于严格．在改良的修正重组中，交付义务的范围可以扩大到债务重组后的60个月之内到期的债券．这种重组在欧洲比较流行．

4. 出现信用事件时会发生什么?

保护的卖方可以选择回收价值的类型．一般有两种交割方式，分别是实体交割和现金交割．实体交割中，违约保护的卖方以面值购买买方所持债券．买方拥有选取最便宜可交割债券的权力．零息债券的名义金额按照承诺价值进行调整．现金交割中，卖方根据回收率以现金形式补偿债券面值与违约事件发生后债券价值之间的差额．

5. 什么是 CDS 溢价?

违约保护的买方支付给卖方的保费通常又称作溢价，量化为合约名义金额每年的基点，通常按季度支付．这种溢价与企业债券或政府债券的溢价概念不同．CDS溢价是年保费占名义金额的百分比，并非基于无风险债券或者任何基准利率．

支付定期的保费之后，如果发生违约，保护的买入方可以以债券面额交付债券，或者获得债券面额与回收价值的差额．因此，CDS类似于企业债券的看跌期权，使保护的买入方免于信用事件使债券价值下跌而造成的损失．也就是说，CDS可以看作看跌期权的佣金，而佣金的支付是在合约期间分期进行的．

举个例子，根据彭博，2013年4月22日摩根斯坦利5年期CDS溢价为150基点．这意味着如果某人以摩根斯坦利信用为参考信用，为200万美元购买5年期的信用保护，那么他每季度需要支付$\frac{150}{4}=37.5$个基点的保费，也就是$\frac{20\ 000\ 000\times 37.5}{10\ 000}=75\ 000$美元，以此获得信用保护．

6. CDS 与债券市场有什么联系?

理论上，CDS溢价与债券收益溢价或者无风险政府债券的超额收益有着密切的联系．为了明白这一点，考虑以下投资组合：

(1) 某公司CDS的空头(即保护的卖方).

(2) 无风险债券的多头．

376

再考虑另一个投资组合：简单持有该公司债券，其期限都相同且名义金额都为100美元．这两笔投资的回报应该完全相同，因此CDS溢价等于企业债券溢价．

如果没有发生违约，到期日CDS和无风险债券的投资组合的本金支付为100美元，因为

CDS 空头方无需支出，而无风险债券得到 100 美元．公司债券的收益也是 100 美元，因为没有发生违约．另一方面，如果发生违约，投资组合的收益为 100 美元减去依赖于违约的或有性支付．该支付金额取决于违约企业债券的回收率．例如，假定回收率为 40%，那么保护的卖出方必须支付 60 美元，即名义金额 100 美元的 60%．使用同样的回收率，由于违约，企业债券的回报也将为 60 美元．这两项投资有着相同的回报和风险敞口．因此，CDS 和该企业债券应该具有相同的溢价水平．

7. 可违约债券与无风险债券的联系

通过研究零息债券与风险/可违约债券的关系，我们可以找到溢价的表达式．回顾之前的内容，零息债券价格可表示为：

$$P(t,T)=E_t(\mathrm{e}^{-\int_t^T r_u \mathrm{d}u}) \tag{23.1}$$

假定利率为常数 $r_u=r$，可得

$$P(t,T)=\mathrm{e}^{-r(T-t)} \tag{23.2}$$

现在，如果我们假设风险比或信用溢价为常数，那么基于以上条件，那么风险/可违约债券价格为：

$$\widetilde{P}(t,T)=\mathrm{e}^{-(r+s)(T-t)} \tag{23.3}$$

这两种债券的联系如下：我们令 p 为违约概率．如果风险债券发生违约，债券持有者在到期日将收到 R，即回收率，这种情况发生的概率为 p．如果不发生违约，债券持有者在到期日将收到 1，这种情况的概率为 $1-p$．那么债券持有者在到期日收入的期望为 $(1-p)\times 1+p\times R$．因此风险债券的价格等于该期望的折现值，即

$$\widetilde{P}(t,T)=P(t,T)((1-p)\times 1+p\times R) \tag{23.4}$$

或者等价地

$$\mathrm{e}^{-(r+s)(T-t)}=\mathrm{e}^{-r(T-t)}((1-p)+pR) \tag{23.5}$$

$$\mathrm{e}^{-s(T-t)}=1-p(1-R) \tag{23.6}$$

根据该等式，我们可以得到信用溢价，记为 s

$$s=\frac{-\ln(1-p(1-R))}{T} \tag{23.7}$$

有两种可能的情况：

已知违约概率 p，求解溢价 s；或者已知溢价 s，计算违约概率 p．

已知 s，我们可以用以下方法计算 p：

$$p=\frac{1-\mathrm{e}^{-s(T-t)}}{1-R} \tag{23.8}$$

例 6 已知 s 计算 p．

考虑摩根斯坦利 5 年期 CDS，其溢价为 150 基点，[⊖] 假定回收率为 50%．利用方程 (23.8)，可得

⊖ 数据来自彭博．

$$p = \frac{1 - e^{-0.015(5)}}{1 - 0.5} \tag{23.9}$$

377

$$p = 14.45\% \tag{23.10}$$

我们可以将这种联系拓展到浮动利率债券(FRN). 考虑支付 Libor 利率的无违约债券，每100 美元面值的债券售价为 100 美元．我们将它与支付 Libor 利率以及 150 基点的 5 年期 FRN 比较. 在这种情况下，5 年期违约保护的保费应该是 150 基点每年．如果违约互换保费小于 1.5%，那么交易者应该购买该债券、购买违约保护，同时卖出支付 Libor 利率的无违约债券．此时套利利润等于 1.5%减去保费．如果保费高于 1.5%，交易者可以卖空该债券，并出售违约保护，买入无违约债券，那么套利利润等于保费减去 1.5%.

23.1.2 担保债务凭证

担保债务凭证(CDO)是一种结构化的资产担保债券(ABS)，它由不同的等级构成，由特殊目的实体发行，以债券、贷款等债务作抵押．每个等级对应的风险和回报都不同，可以满足不同风险偏好的投资者．CDO 的价值和收益是根据固定收益标的资产所构成的投资组合得到的．CDO 中的不同等级对应着不同的风险类型，其中高等级的部分是安全系数最高的，一旦发生损失，首先要保障高等级的本金和利息．因此低等级部分的回报率更高，或者发行价格更低，以此来平衡它们所承担的违约风险．

1. 现金 CDO

现金 CDO 是现金资产的投资组合，包括贷款、高收益债券，如企业债券、资产担保债券(ABS)或者抵押担保债券(MBS). 此外，有些 CDO 结构要求相关人员(投资组合经理)对投资组合进行积极管理，从而维持资产基础、对将到期的基金再投资．基于这个原因，资产的所有权被转移到了法人实体，即特殊目的机构(SPV)，专门负责发行 CDO 份额．资产遭受损失的风险从高到低分配到从低到高的 CDO 等级．现金 CDO 的发行量在 2012 年超过了 8000 亿美元．[1]

2. 合成 CDO

合成 CDO(synthetic CDO)不含有现金资产如债券和贷款．合成 CDO 的风险敞口是具有固定收益的资产组合，但它并不拥有这些资产，而是通过卖出相应的 CDS 保护来承担风险．CDO 作为信用保护的卖方，将收到定期的现金支付，也就是保费．作为交换，将在相应资产发生违约或者信用事件时承担一定损失．此外，和现金 CDO 类似，构成合成 CDO 时将形成不同的等级，各等级对风险的承担情况不同．损失首先由“股权等级”承担，然后是中间等级，最后是高级等级．每个等级都会定期收到支付(即保费)，等级越低的份额获得的保费越高．

CDO 份额可能需要融资，也可能不需要．由于签订了互换协议，合成 CDO 只有在发生一定的信用事件时，才会发生支付行为．一部分的投资者是需要投入资金的．尤其是低

---

[1] 数据来自彭博．

级份额投资者，它们承受的风险最大，需要投入资金作为抵押．在发生信用事件之前，投入的资金通常会用于高质量流动性资产的投资，或者用于担保投资协议(GIC)，从而获得略低于 Libor 利率的收益．这些投资的收益以及保费构成了各等级的利润．当信用事件发生时，需要对保护的买方支付补偿，这笔补偿就来自 GIC 或者流动性投资．相比之下，高等级通常不需要投资，因为他们的风险要低得多．与现金 CDO 不同，合成 CDO 的高等级投资者能够定期地收到回报，却不需要在开始时做任何投资．在信用事件发生后，如果投资组合的损失达到了高级份额，那么这部分投资者需要进行支付．合成 CDO 的发行量在 2012 年超过了 1200 亿美元(数据来自彭博)．从发行的角度看，合成 CDO 所需时间更短．不需要购买或管理现金资产，且 CDO 等级结构更为明确．

378

3. CDO 市场的特征

现金 CDO 包含企业债券池或者结构性金融资产池，如住宅抵押贷款证券(RMBS)、商业房产抵押贷款支持证券(CMBS)以及资产抵押证券(ABS)．合成 CDO 由一个很大的 CDS 池组成(通常超过 100 种)．合成 CDO 近年来越来越流行，尤其是在欧洲市场，超过 90%的 CDO 交易是合成 CDO. 在美国，合成 CDO 的交易量占所有套利型 CDO 的三分之一．合成 CDO 的结构比现金 CDO 更加灵活．

23.2 信用违约互换的定价

信用违约互换的定价主要有两种方法．第一种叫做*结构法*，第二种叫做*简化法*．顾名思义，结构法假设当债务人[⊖]的市场价值下降到一定的边界值时，违约会发生．而简化法将违约事件看作随机(无法预测的)事件直接进行建模．

在对信用违约互换定价前，我们需要用参数对以下概念进行量化：

债券违约的可能性、违约的概率；

违约发生后的回收率；

信用事件对流动性、规律性、市场情绪的影响．

在对 CDS 定价时，通常使用 Libor 利率作为无风险折现率．定价理论表明，一份衍生品的价格等于使用其他证券构建无风险投资组合来复制该衍生品所需的费用．由于大多数信用衍生品的市场交易者是银行，它们的资金利率接近 Libor，因此吸收这些投资的成本也应该是 Libor. 所以，在对 CDS 以及其他信用衍生品定价时，Libor 利率是有效的无风险利率．

在信用建模时，结构法和简化法是互斥的．本节我们将介绍 CDS 定价的一般设定，以及简化法的定价过程．我们还将在 23.4 节介绍一些一般的结构法．

⊖ 债务人即借债者，可以是人也可以是企业．他们欠债的对象叫债权人，也可以是人或企业．其债务关系来自于借贷或发行债券．

23.2.1 一般设定

在典型的 CDS 合约中，我们通常有两支潜在的现金流：一支是固定性的保费，另一支是“或有性”的补偿．前者是指，保护的买入方支付定期、固定的 CDS 保费，直到合约到期或者发生违约．后者是指，当且仅当违约发生，保护的卖出方向买方支付一笔补偿．或有性支付的金额一般等于名义金额乘以$(1-R)$，其中 R 是回收率．表示关于名义金额的百分比．

379

对于保护买入方而言，在任意一个时间点，CDS 合约的价值等于或有性现金流的现值减去需要付出的保费的现值，即

CDS 价值(对保护买入方而言)＝PV[或有性现金流]－PV[固定性(保费)现金流]

要计算这些现值，我们需要知道违约概率，即标的信用的信用曲线，还有发生违约的回收率、无风险折现因子、收益率曲线．此外交易对手风险也是一个影响较小的因素．为简化起见，我们假设不存在交易对手风险，互换的名义价值为 1000 万美元．

1. 对定期付款的保费估值

首先我们来看固定性现金流．在每个付款日，定期支付额按照年 CDS 保费 s 乘以累积天数所对应的 Δ_t 计算．例如，如果 CDS 保费为每年 200 基点，按季度支付，那么每次支付$\Delta_t s=\frac{200}{4}=50$基点．但是该付款只有当参考信用在付款日前没有违约的情况下才存在．因此我们必须考虑生存概率，即参考信用在支付日之前没有发生违约的概率．例如，如果参考信用前 3 个月的存活率是 95%，那么在 3 个月后的 t_1 时刻，期望支付额为 $q(t_1)\Delta_1 s=\frac{0.95\times200}{4}=47.5$ 基点，其中 $q(t)$ 为 t 时刻的存活率．然后，使用支付日对应的折现因子$P(t_0,t_i)$，即可求得付款额的现值 $P(t_0,t_i)q(t_i)\Delta_t s$. 将每次付款的现值加起来，可得

$$\sum_{i=1}^{n}P(t_0,t_i)q(t_i)s\Delta i \tag{23.11}$$

如果我们将以上表达式中的溢价 s 去掉，就可以得到总支付额中一个基点所对应的现值，条件是参考实体不发生违约．这在 CDS 定价中是一个重要的量，通常叫做风险 PV01，即风险单位基点的现值．

2. 对应计付款保费的估值

为了更为精确，我们还需要考虑如果定期付款日之间发生了违约事件，那么在此之前应付的保费是多少．我们假设违约发生在两个付款日的中间．也就是说假如参考实体在付款日 t_{i-1} 和 t_i 之间违约，则这段时间的应计付款额是$\frac{\Delta_i}{2}s$. 我们需要用这个值乘以在该事件区间发生违约的概率．也就是参考信用维持到付款日 t_{i-1}，但没有维持到付款日 t_i 的概率．这个概率可以写为 $q(t_{i-1})-q(t_i)$. 因此，在某一特定区间，应计付款的期望值为

$(q(t_{i-1})-q(t_i))\frac{\Delta_i}{2}s$. 因此，所有应计付款期望值的现值为

$$\sum_{i=1}^{n}P(t_0,t_i)(q(t_{i-1})-q(t_i))s\frac{\Delta_i}{2} \tag{23.12}$$

380

将这两部分保费相加，即可得到固定付款额的现值

$$\sum_{i=1}^{n}P(t_0,t_i)q(t_i)s\Delta_i+\sum_{i=1}^{n}P(t_0,t_i)(q(t_{i-1})-q(t_i))s\frac{\Delta_i}{2} \tag{23.13}$$

3. 对或有性补偿的估值

下面我们计算或有性补偿的现值．假设参考实体在付款日 t_{i-1} 和 t_i 之间违约．保护的购买方将收到$(1-R)$的补偿．这笔付款当且仅当参考信用违约后才存在，我们需要将它乘以 $q(t_{i-1})-q(t_i)$，即违约发生在这段时间区间的概率．将每个区间内的期望补偿额折现，即可得到或有性补偿的现值

$$(1-R)\sum_{i=1}^{n}P(t_0,t_i)(q(t_{i-1})-q(t_i)) \tag{23.14}$$

4. CDS 定价公式

确定了定期保费和或有性补偿之后，我们可以得到计算 CDS 价值的公式．

$$\begin{aligned}\mathrm{PV(CDS)}=&\sum_{i=1}^{n}P(t_0,t_i)q(t_i)s\Delta_i+\sum_{i=1}^{n}P(t_0,t_i)(q(t_{i-1})-q(t_i))s\frac{\Delta_i}{2}\\&-(1-R)\times\sum_{i=1}^{n}P(t_0,t_i)(q(t_{i-1})-q(t_i))\end{aligned}$$

当双方进入 CDS 交易时，可以根据互换交易的价值为 0，从而确定 CDS 溢价(定期保费的价值等于或有性补偿的价值). 因此，存在以下等式：

$$\begin{aligned}&\sum_{i=1}^{n}P(t_0,t_i)q(t_i)s\Delta_i+\sum_{i=1}^{n}P(t_0,t_i)(q(t_{i-1})-q(t_i))s\frac{\Delta_i}{2}\\&=(1-R)\sum_{i=1}^{n}P(t_0,t_i)(q(t_{i-1})-q(t_i))\end{aligned} \tag{23.15}$$

已知所有参数，保费年金 s 为：

$$s=\frac{(1-R)\sum_{i=1}^{n}P(t_0,t_i)(q(t_{i-1})-q(t_i))}{\sum_{i=1}^{n}P(t_0,t_i)q(t_i)\Delta_i+\sum_{i=1}^{n}P(t_0,t_i)(q(t_{i-1})-q(t_i))\frac{\Delta_i}{2}} \tag{23.16}$$

例 7 假定对以下 CDS 交易定价．

考虑一个按季度支付保费的 5 年期 CDS. 我们假定信用溢价为 180 基点，回收率为 50%，名义总额为 1000 万美元．折现曲线和各付款日期的存活率如表 23-1 所示．

表 23-1 折现曲线和存活概率

时间/月	折现曲线 $P(t_0, t_i)$	直到 t_i 的存活概率/%
0	1.000 00	100.0
3	0.999 38	99.9
6	0.998 75	99.7
9	0.998 13	99.5
12	0.997 50	99.2
15	0.996 26	98.9
18	0.995 01	98.4
21	0.993 77	97.9
24	0.992 53	97.2
27	0.990 67	96.5
30	0.988 81	95.5
33	0.986 96	94.5
36	0.985 11	93.0
39	0.982 65	91.5
42	0.980 20	89.5
45	0.977 75	87.5
48	0.975 31	85.5
51	0.971 66	82.5
54	0.968 02	79.5
57	0.964 40	76.5
60	0.960 79	73.5

根据表 23-1 中的值我们可以计算每个时间段固定付款的现值．已知生存概率，我们可以通过 $q(t_{i-1})-q(t_i)$ 计算违约概率，并以此计算应计付款的现值．此外根据违约概率我们可以计算每个时间段或有性付款的现值． ◀

所有期望固定付款的现值可以通过每一期的付款乘以相应的生存概率，再用无风险利率折现最后对各时期求和得到．例如，9 期即 27 个月的折现因子为 0.990 67，生存概率为 96.5%，因此固定性付款的现值为

$$P(t_0,t_i)q(t_i)\frac{s}{4}N = 0.990\ 67\times\frac{96.5}{100}\times\frac{\frac{200}{4}}{10\ 000}\times 10\ 000\ 000 \tag{23.17}$$

$$= 47\ 800(\text{美元}) \tag{23.18}$$

381

如表 23-2 所示．表 23-2 最后一行的 905 447 美元是所有此类付款现值的和．如果在两个付款日 t_{i-1} 和 t_i 之间发生违约，我们假定它发生在该时间段的中间，则应计保费额为 $\frac{200}{4}=50$ 基点的一半，即 25 基点．该时段应计付款的期望值等于 25 基点乘以违约概率 $q(t_{i-1})-q(t_i)$，再折现后即可得到现值．

$$P(t_0,t_i)(q(t_{i-1})-q(t_i))\frac{\frac{S}{4}}{2}N = 0.990\ 67\times\frac{97.2-96.5}{100}$$

$$\times \frac{\frac{200}{4}}{2 \times 10\ 000} \times 10\ 000\ 000 \tag{23.19}$$

382

$$= 0.990\ 67 \times \frac{0.7}{100} \times \frac{25}{10\ 000} \times 10\ 000\ 000 \tag{23.20}$$

$$= 173(\text{美元}) \tag{23.21}$$

表 23-2 定期付款和或有性付款的现值的现金流

时间/月	折现因子 $P(t_0,\ t_i)$	到 t_i 的存活概率/%	定期付款的现值/美元	Δ_i 期间违约的概率/%	应计付款的现值/美元	或有性付款的现值/美元
0	1.000 00	100		0	0	0
3	0.999 38	99.9	49 919	0.1	25	4997
6	0.998 75	99.7	49 788	0.2	50	9988
9	0.998 13	99.5	49 657	0.2	50	9981
12	0.997 50	99.2	49 476	0.3	75	14 963
15	0.996 26	98.9	49 265	0.3	75	14 944
18	0.995 01	98.4	48 955	0.5	124	24 875
21	0.993 77	97.9	48 645	0.5	124	24 844
24	0.992 53	97.2	48 237	0.7	174	34 738
27	0.990 67	96.5	47 800	0.7	173	34 673
30	0.988 81	95.5	47 216	1.0	247	49 441
33	0.986 96	94.5	46 634	1.0	247	49 348
36	0.985 11	93.0	45 808	1.5	369	73 883
39	0.982 65	91.5	44 956	1.5	368	73 699
42	0.980 20	89.5	43 864	2.0	490	98 020
45	0.977 75	87.5	42 777	2.0	489	97 775
48	0.975 31	85.5	41 694	2.0	488	97 531
51	0.971 66	82.5	40 081	3.0	729	145 749
54	0.968 02	79.5	38 479	3.0	726	145 203
57	0.964 40	76.5	36 888	3.0	723	144 660
60	0.960 79	73.5	35 309	3.0	721	144 118
			905 447 美元		6467 美元	1 293 431 美元

将所有时段内的值折现并求和，即可得到所有应计固定付款的期望现值．表 23-2 最后一行的金额 6467 美元是所有这些付款现值的和．不出所料，这是一个很小的值，因为这是各时期违约概率与应计付款的乘积求和得到的，这两个都是较小的值，单独时段的违约概率较低且发生违约后该时段的应计付款为 25 基点．根据以上内容，我们可以得到 5 年期 1000 万美元所对应的固定付款现值，或者说保护购买方的期望支出的现值．

某时段或有性补偿的期望值等于该时段违约发生的概率乘以$(1-R)$乘以名义金额．回收率为 50%，我们将 0.5 乘以各时段违约概率乘以名义金额，得到期望值，然后折现．

383

$$P(t_0,t_i)(q(t_{i-1})-q(t_i))(1-R)N$$

$$=0.99067\times\frac{97.2-96.5}{100}\times(1-0.50)\times 10\,000\,000 \tag{23.22}$$

$$=0.99067\times\frac{0.7}{100}\times 0.50\times 10\,000\,000 \tag{23.23}$$

$$=34\,673(\text{美元}) \tag{23.24}$$

将整个互换协议期限内各阶段的值加起来，即可得到或有性补偿的期望现值，为 1 293 431 美元，见表 23-2 的最后一行．

因此，已知该 5 年期 CDS 的溢价为 200 基点/年，对于信用保护买入方(固定保费支付方)而言，该 CDS 的价值为

$$\text{CDS 价值}=\text{或有性付款的现值}-\text{定期付款的现值} \tag{23.25}$$

$$=1\,293\,431-(905\,447+6467) \tag{23.26}$$

$$=381\,517(\text{美元}) \tag{23.27}$$

直观地说，该 CDS 期限内平均每年的违约概率为 5.3%，因为信用维持 5 年的概率为 73.5%．由于回收率为 50%，每年的平均期望损失为(1－0.5)×5.3%＝2.65%．CDS 溢价为 200 基点每年，表示本例中信用保护买入方可以支付 200 基点来获得信用保护，而信用风险造成的期望损失为每年 265 基点．毫无意外，这份交易对于 CDS 买入方而言是价值大于零的，对 1000 万美元的名义金额而言，这份 CDS 的价值为 381 517 美元，也就是约 381 基点．

5. 信用违约互换头寸的估值

假设初始时刻为 0，t 时刻一份信用违约互换的市场价等于被保护的价值与保费的成本之差．

$$V(t)=\pm(\text{保护的价值}-\text{保费的成本}) \tag{23.28}$$

假定 t 时刻市场隐含溢价为 $s(t)$，则该互换的价值等于初始溢价与当前溢价的差值乘以风险 PV01，即

$$V_t=(s(t)-s(0))\text{PV01} \tag{23.29}$$

这里 PV01 的定义前面介绍过，表示在 CDS 合约中的付款日所付保费的现值的一个基点．

在利率和信用曲线为水平线的假设下，以上计算是较为准确的，尽管作了一定简化．

例 8 某投资者为 250 万名义金额购买了 5 年期的保护，溢价为 100 基点．合约规定保费按年支付．一年后该信用下降，溢价为 200 基点．假设 $r=0.25\%$ 且每年的存活率如表所示，则该 CDS 头寸的价值可计算为：

t	$q(t)$
1	0.97
2	0.94
3	0.91
4	0.89
5	0.86

$$V(t)=N(s(t)-s(0))\text{PV01} \tag{23.30}$$

$$V(t)=N\frac{(200-100)}{10\,000}\text{PV01} \tag{23.31}$$

$$\text{PV01}=\sum_{i=1}^{n}P(t_0,t_i)q(t_i)\Delta_i \tag{23.32}$$

$$\text{PV01} = 4.5403 \tag{23.33}$$

384

$$V(t) = N\frac{(200-100)}{10\,000}4.5403 = 1.1351 \tag{23.34}$$ ◀

23.2.2 简化法——风险率法

上一节我们介绍了在已知各付款日期间的存活和违约概率的条件下，如何对 CDS 定价. 本节我们将介绍简化法，该方法依赖于风险率，需要对概率建模. 在介绍风险率方法之前，我们首先定义风险率，并介绍其计算方法.

1. 风险率的定义与计算

定义 $q(t)\Delta t$ 为 t 到 $t+\Delta t$ 时刻之间的违约概率，观察时刻为 0 时刻. 风险率 $h(t)$ 相当于违约密度，它使得 $h(t)\Delta t$ 表示 t 到 $t+\Delta t$ 时刻之间发生违约的概率，观察时刻为 t，同时假设 0 到 t 之间不发生违约. $q(t)$ 和 $h(t)$ 之间有以下关系：

$$q(t) = h(t)\mathrm{e}^{-\int_0^t h(\tau)\mathrm{d}\tau} \tag{23.35}$$

我们将逐步证明该方程. 设 $F(t)$ 为 0 到 t 时刻之间发生违约的概率，则

$$F(t) = \int_0^t q(u)\mathrm{d}u \tag{23.36}$$

直觉上，我们可以说 0 到 $t+\Delta t$ 时刻之间发生违约的概率，等于 0 到 t 发生违约的概率加上存活到 t 并在 t 到 $t+\Delta t$ 之间违约的概率. 用公式表示，即为

$$F(t+\Delta t) = F(t) + (1-F(t))h(t)\Delta t \tag{23.37}$$

我们可以从中得到 $h(t)$ 的表达式

$$h(t) = \frac{1}{1-F(t)}\frac{F(t+\Delta t)-F(t)}{\Delta t} \tag{23.38}$$

$$= \frac{1}{1-F(t)}\frac{\Delta F}{\Delta t} \tag{23.39}$$

令 Δt 趋于 0，取极限可得

$$h(t) = \frac{1}{1-F(t)}\frac{\mathrm{d}F}{\mathrm{d}t} \tag{23.40}$$

由于生存是违约的对立面，我们有 $S(t)=1-F(t)$. 将其带入方程(23.40)可得

$$\frac{\mathrm{d}S(t)}{\mathrm{d}t} = -S(t)h(t) \tag{23.41}$$

初始条件为 $S(0)=1$. 求解该常微分方程，可得

$$\ln S(t) = -\int_0^t h(u)\mathrm{d}u + c \tag{23.42}$$

利用初始条件可得 $c=0$，因此有

$$S(t) = \exp\left(-\int_0^t h(u)\mathrm{d}u\right) \tag{23.43}$$

用 $t+\mathrm{d}t$ 时刻的生存概率减去 t 时刻的生存概率，即可得到 t 到 $t+\mathrm{d}t$ 之间发生违约的概率，

观察时刻为 0 时刻，即

$$S(t+\mathrm{d}t)-S(t)=q(t)\mathrm{d}t \tag{23.44}$$

或

$$q(t)=\frac{S(t+\mathrm{d}t)-S(t)}{\mathrm{d}t}=\frac{\mathrm{d}S}{\mathrm{d}t} \tag{23.45}$$

我们可以通过对(23.43)求导计算$\frac{\mathrm{d}S}{\mathrm{d}t}$，由于$\frac{\mathrm{d}S}{\mathrm{d}t}$等于$q(t)$，可得

$$q(t)=h(t)\mathrm{e}^{-\int_0^t h(\tau)\mathrm{d}\tau} \tag{23.46}$$

风险率法是应用最为广泛的结构法．该方法首先由 Jarrow 和 Turnbull 提出，运用于受信用风险影响的金融衍生品定价中(Jarrow 和 Turnbull，1995)．他们将信用事件定义为泊松计数过程的首次事件，t 时刻发生的概率为

$$P(\tau\leqslant t+\mathrm{d}t \mid \tau>t)=h(t)\mathrm{d}t \tag{23.47}$$

385

$h(t)$就是风险率．我们已经证明生存概率可表示为

$$E^Q\left(\exp\left(-\int_0^T h(u)\mathrm{d}u\right)\right) \tag{23.48}$$

在这种设定下，很明显风险率是确定性的，且与利率或回收率相独立．[⊖]

定义 $\lambda_t=1_{\tau\leqslant t}$. 可证明 $\lambda_t-\int_0^t(1-\lambda_u)h_u\mathrm{d}u$ 是鞅．λ_t 的取值为 0 或 1. $\lambda_t=0$ 意味着对于所有 $u<t$ 有 $\lambda_u=0$,因此 $\lambda_t-\int_0^t(1-\lambda_u)h_u\mathrm{d}u=-\int_0^t h_u\mathrm{d}u$. 若 $\lambda_t=1$,则存在 t^* 介于 0 和 t 之间,使得对于 $u\in(0,t^*)$ 有 $\lambda_u=0$. 因此 $\lambda_t-\int_0^t(1-\lambda_u)h_u\mathrm{d}u=1-\int_0^t h_u\mathrm{d}u$.

以下随机指数过程是鞅

$$\varepsilon\left(-\lambda_t+\int_0^t(1-\lambda_u)h_u\mathrm{d}u\right) \tag{23.49}$$

$$\varepsilon\left(-\lambda_t+\int_0^t(1-\lambda_u)h_u\mathrm{d}u\right)=\mathrm{e}^{\int_0^t(1-\lambda_u)h_u\mathrm{d}u}\Pi(1-\Delta\lambda) \tag{23.50}$$

因此

$$E(\mathrm{e}^{\int_0^t(1-\lambda_u)h_u\mathrm{d}u}(1-\lambda_t))=1 \tag{23.51}$$

这里$(1-\lambda_u)$是多余的，因此

$$E(1-\lambda_t)=\mathrm{e}^{-\int_0^t h_u\mathrm{d}u} \tag{23.52}$$

根据信用溢价，我们可以利用迭代过程求违约概率．在等价回收模型(Jarrow 和 Turnbull，1995)中，回收率为外生的常数，我们可以假设无风险利率 $r(t)$和溢价 $s(t)$的变化都服从 Longstaff 和 Schwartz 框架(Longstaff 和 Schwartz，1992)，正如 Hatgioannides

⊖ 信用投资组合模型通常假设回收率与违约概率是独立的．John Frye(2003)通过充足的证据表明，根据历史违约数据，这种假设是不成立的．我们会在后面强调这一点．

和 Petropoulos(2007)所介绍．

此外，对于一小段时间 dt，短期信用差价直接与局部违约概率 $q(t)$有关．局部违约概率是 t 到 $t+dt$ 之间的违约概率，条件是 t 时刻以前没有违约发生．此外，我们可以将违约概率与违约过程的密度联系起来(Arvanitis 等，1999).

现在利用基于溢价的模型得到溢价折现因子，推导局部违约概率，进而推导条件违约概率．随着局部违约概率的确定，我们可以进一步确定累积生存概率，并在每个时间点将违约和不违约的概率都考虑进来．通过 Jarrow，Lando 和 Turnbull 的评分模型，我们可以估计风险溢价，我们可以利用 Longstaff 和 Schwartz(1992)模型得到隐含的评分转移矩阵．

2. 常数风险率模型

关于信用违约互换的盈亏平衡溢价，一个很自然的想法是它应该使得保费的现值等于保护的现值．在计算现值时，违约事件的发生时刻起着很关键的作用．我们可以考虑违约发生在每个区间[t，$t+dt$]的情况，然后积分．也就是说我们假定 t 之前存活，[t，$t+dt$]之间违约．生存到 t 时刻的概率为 $q(t)$，故上述概率为 $q(t)h(t)dt$，而违约后的补偿为(1－R)，我们需要用无风险利率折现．假定无风险利率为常数，则我们可以将保护的现值写为

386

$$(1-R)\int_0^t h\exp(-(r+h)u)\mathrm{d}u=\frac{h(1-R)(1-\mathrm{e}^{-(r+h)t})}{r+h}\tag{23.53}$$

保费的价值是期限内违约前每季度付款的现值．假定溢价为常数，保费的现值为

$$s\int_0^T \mathrm{e}^{-(r+h)u}\mathrm{d}u=\frac{s(1-\mathrm{e}^{-(r+h)T})}{r+h}\tag{23.54}$$

令这两个方程相等，可以求解出 s

$$s=h(1-R)\tag{23.55}$$

以上关系叫作信用三角形，因为只要知道其中两个量，就可以计算出第三个．以上公式可以理解为，在任何一个小时段内，投资者支付的保费都恰好补偿他面临的违约风险．

例 9 信用三角公式

某公司信用溢价为 200 基点，若我们假设回收率为 50%，根据该公式，我们可以计算出隐含风险率为 4% $\left(h=\frac{0.02}{0.5}=0.04\right)$. 利用方程(23.55)，我们可以求出 5 年的生存概率曲线．◀

3. bootstrap 风险率

在假设利率和信用曲线都是水平线的情况下，以上计算是较为准确的，尽管作了一定简化．现在我们放宽风险率为常数这一假设，利用 bootstrap 构造风险率的期限结构．我们可以通过以下步骤完成自助抽样．

我们先从期限最短的 CDS 入手，找到能够使两支付款的期望现值相等的风险率．然后我们拓展到期限第二短的 CDS，并假设风险率是分段常数．也就是说，我们假设第一段期限内风险率为 h_1，在 t_1 和 t_2 之间风险率为 h_2. 固定 h_1，我们可以用之前令现值相等的方法，求出 h_2. 然后继续此过程，直到期限达到最长．

假设每 6 个月付款一次，第一个 CDS 的溢价为 s_1，我们有

$$(1-q_1)\frac{s_1}{2}Ne^{-L_1t_1}=q_1N(1-R)e^{-L_1t_1} \tag{23.56}$$

求解 q_1 可得

$$q_1=\frac{\frac{1}{2}s_1\mathrm{d}f_1}{\left((1-R)+\frac{s_1}{2}\right)\mathrm{d}f_1} \tag{23.57}$$

设第二个 CDS 的溢价为 s_2，我们有

$$\begin{aligned}&(1-q_1)\frac{s_2}{2}Ne^{-L_1t_1}+(1-q_2)\frac{s_2}{2}Ne^{-L_2t_2}\\&=q_1N(1-R)e^{-L_1t_1}+(q_2-q_1)N(1-R)e^{-L_2t_2}\end{aligned} \tag{23.58}$$

求解 q_2 可得

$$q_2=\frac{\frac{1}{2}s_2\mathrm{d}f_2+(1-q_1)\frac{s_2}{2}\mathrm{d}f_1-(1-R)q_1(\mathrm{d}f_1-\mathrm{d}f_2)}{\left((1-R)+\frac{1}{2}s_2\right)\mathrm{d}f_2} \tag{23.59}$$

根据归纳法我们可以证明

$$q_k=\frac{\frac{1}{2}s_k\mathrm{d}f_k+\sum_{i=1}^{k-1}\left((1-q_i)\frac{s_k}{2}\mathrm{d}f_i-(1-R)q_i(\mathrm{d}f_i-\mathrm{d}f_{i+1})\right)}{\left((1-R)+\frac{1}{2}s_k\right)\mathrm{d}f_k} \tag{23.60}$$

在信用衍生品定价中，回收率是风险中性世界下的期望回收率，信息来源仅有价格信息. 然而，在已知价格的情况下，很难求出回收率，因为它还与违约概率有着复杂的关系. 市场参与者一般采用的方法是利用信用评分机构的违约研究估计回收率，可能通过代理人来完成这一工作. 这种评分是点对点的，代理人需要根据美国和非美国实体的数据进行调整. 用这种方法估测回收率有一个缺陷，这些结果是通过回顾过去得到的，而且只计入了违约，没有考虑重组. 此外，研究显示回收率与违约呈负相关，为了反映这种关系，研究者引入了随机回收率模型. 对回收率建模的一般方法是使用 Beta 分布.

387

现实世界违约概率比定价时计算的要低得多. 这与期权的隐含波动率和历史价格表现出的实际波动率的关系类似. 这是因为人们对市场风险有厌恶情绪，引起了供需不平衡，使得风险保费偏高.

23.3 多家公司信用产品的定价

正如合成 CDO 一样，许多信用产品依赖于不止一个参考实体. 因此，对于这类产品，我们不但需要对投资组合内的每个参考实体的违约概率建模，还要考虑所有资产之间的相关性. 在下一节我们将会看到，通常我们用含 Copula 的结构模型来研究该问题，其中 Copula 函数的作用是描述变量之间的相关性.

23.3.1 违约相关性建模

本节我们将介绍如何对信用衍生品定价中的联合违约概率建模．为介绍这些内容，首先我们要拓展结构法和简化法，从而适应多个参考实体的问题．原则上说这种拓展对两种方法都是可行的，但实际上，结构法在实践者中的使用更为广泛，因为该方法在高维情况下更容易处理．在介绍这些拓展之前，我们先学习一些估测联合违约时所需的基础概率概念．有许多不同的模型能为这些衍生品定价．

1. 联合违约概率的估计

考虑两个债务人 A 和 B，以及某个具体的界限 T. 令

- p_i 为债务人 i 在 T 之前违约的概率．
- p_{ij} 为 i、j 都在 T 之前违约的概率，即联合违约概率．
- $p_{i|j}$ 为已知 j 在 T 之前违约，i 也在 T 之前违约的概率．

由概率论的知识，我们知道

$$p_{A|B}=\frac{p_{AB}}{p_B} \tag{23.61}$$

$$p_{B|A}=\frac{p_{AB}}{p_A} \tag{23.62}$$

$$\rho_{AB}=\text{线性相关系数或违约事件相关系数} \tag{23.63}$$

$$=\frac{p_{AB}-p_Ap_B}{\sqrt{p_A(1-p_A)p_B(1-p_B)}} \tag{23.64}$$

由于违约概率非常小，相关系数 ρ_{AB} 对头寸的联合违约风险影响非常大．

必须注意到该系数的上下限不是－1 和 1，而是边际概率本身的函数．由于缺少违约数据，违约概率的估计非常困难．因此在对联合违约建模时，实践者寻找了替代方法来得到联合违约的频率，以此来避免直接估计联合概率，如拟合法或半解析法．

388

2. 联合违约的结构法

在文献中有多种不同方法可以对有相关性的违约建模，并对多资产的信用衍生品定价．Hull 和 White 的模型通过扩散项相关的资产价格模型并校正违约阈值，能够生成相互相关的违约时间，使得与已知的边际违约概率一致．大多数此类多资产模型需要拟合．少部分如 Finger、Gregory 和 Laurent 所研究的模型不需要拟合，他们的方法以低维因子结构和条件独立性为基础，是半解析的．

我们在拟合过程中生成样本路径．当接收到信用事件的发生情况时，我们可以生成违约路径以及各样本路径上违约发生的确切时间．这是根据先验已知的联合违约分布，对篮子内的债务人进行随机选择得到的．和其他拟合一样，我们可以以此来求多资产信用工具的具体价值．

诸如综合损失份额或第 n 个违约这样形式的多资产信用违约支付不能用单一资产信用

违约的交易的简单组合来替代．我们已经说过市场参与者在对具有相关性的产品估值时，一般采取标准无套利原则．因此在对这些工具估值时，需要考虑所有违约可能性，并计算风险中性期望．

例 10 联合违约

在结构法中，当资产价值下降到某个阈值以下后违约发生．同样的道理，当资产价值的联合动态变化达到一定界限后，联合违约发生．

例如，假设每个资产服从以下动态变化：

$$\ln(S_t^{(i)}/S_0^{(i)}) = \left(r - q - \frac{\sigma^2}{2}\right)t + \sigma W_t^{(i)} \tag{23.65}$$

$$= \left(r - q - \frac{\sigma^2}{2}\right)t + \sigma\sqrt{t}Z_i \tag{23.66}$$

其中 Z_i 是标准正态随机变量．假设对所有债务人，无风险利率为 2%，无红利，波动率为 30%，初始价格 $S_0=100$，期限为 1 年，$t=1$. 同时假定 $\ln(S_t/S_0)<L=-0.3$ 则发生违约事件．考虑 5 种不同的相关系数：0%，25%，50%，75%和 90%. 为了使对比公正，对于每个相关系数我们都取同样的样本，且对每个债务人的样本总价值为 10 000 美元(见图 23-2).

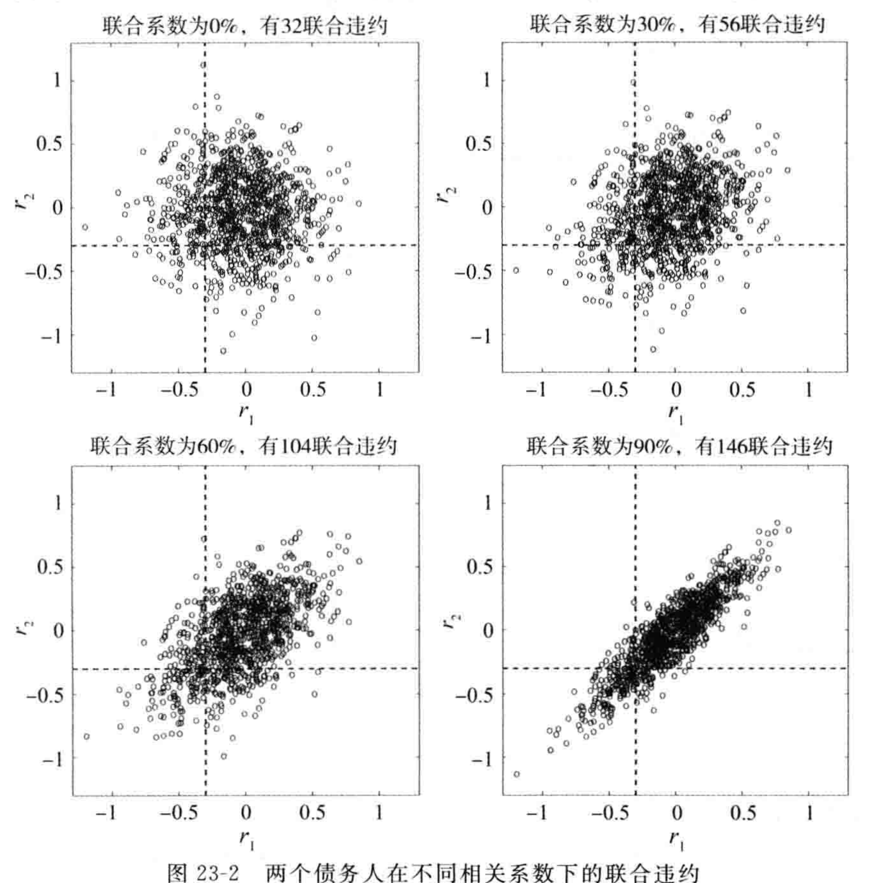


图 23-2 两个债务人在不同相关系数下的联合违约

根据例 10 和图 23-3 很容易得到以下结论，随着资产回报相关性增加，联合违约的概率也增加．这说明资产相关性能够导致违约相关性． ◀

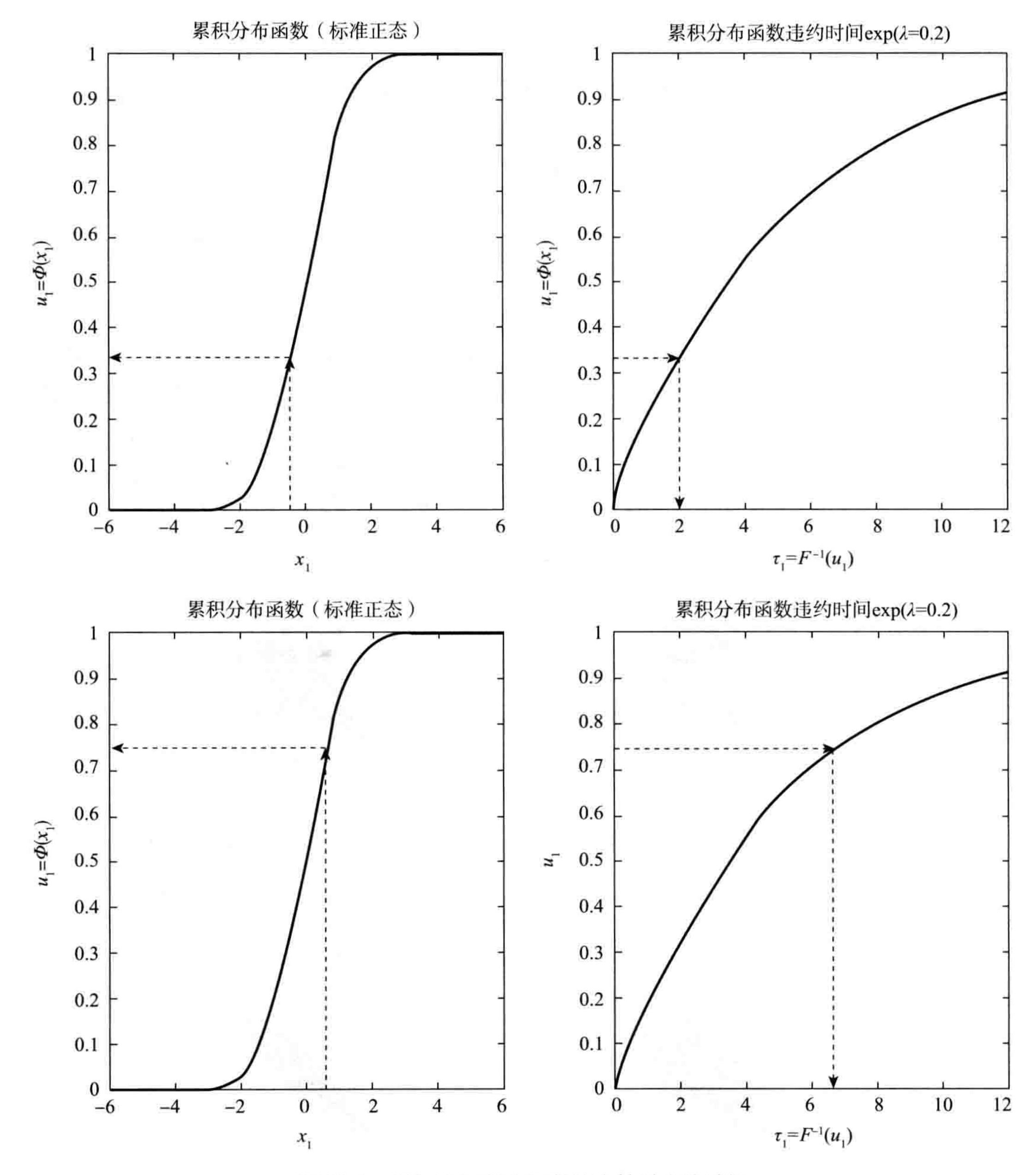


图 23-3 把标准正态随机变量映射到违约时间

3. 高斯 Copula 模型

David Li 的高斯 Copula 模型提供了一种计算量低且简单的拟合相关性违约的方法．它

的基础假设是违约时间的多变量分布和资产回报的多变量分布有着相同的相关性结构．在该论文中，作者假设相关性结构是高斯函数．根据该假设，已知相关性结构之后，整个结构都可以完全确定．

假设第 j 个可违约金融工具的边际分布为 F_j，$j=1, 2, \cdots, k$. 在 Copula 模型中，我们可以将这些边际分布放到一个相关性矩阵 Λ 中．由于资产回报无法直接观察，我们将使用利用股权价值的相关性的作为资产相关性的“代表”.

在有两种资产的情况下，我们可以写出违约事件的联合分布函数

$$P(\tau_1 < x_1, \tau_2 < x_2) = \Phi_{2,\Lambda}(\Phi^{-1}F_1(x_1), \Phi^{-1}(F_2(x_2))) \tag{23.67}$$

389 ~ 391

其中 $\Phi_{2,\Lambda}$ 是双变量标准正态分布函数，相关矩阵为 Λ，Φ^{-1} 为标准正态分布函数的反函数．我们可以很容易地将它推广到 k 维的情况．

根据该分布很容易对违约事件进行拟合．假设有 k 名债务人，相关矩阵为 Λ，我们可以进行以下步骤：

使用相关矩阵 Λ 的 Cholesky 分解，拟合多变量正态随机向量 x，x 是 k 维向量．

建立超立方体 $U=(\Phi(x_1), \cdots, \Phi(x_k))$.

$$\tau = (\tau_1, \cdots, \tau_k) = (F_1^{-1}(u_1), \cdots, F_k^{-1}(u_k))$$

很容易验证 τ 包含了已知的边际且具有正态相关性结构，后者是由相关矩阵 Λ 决定的．这里我们举两个例子，一个是一维情况，另一个是三维情况．

例 11 一维的例子

假设违约时间分布是指数分布，$\lambda=0.2$，$\theta=\dfrac{1}{\lambda}=5$. 该假设意味着违约平均 5 年发生一次．取标准正态分布变量 z，计算 $U=\Phi(z)$，然后求 $\tau=F^{-1}(u)$，其中 $F(x)=1-\exp(-\lambda x)$，这意味着 $F^{-1}(u)=-\dfrac{\ln(1-u)}{\lambda}$. 既然我们已经知道如何通过违约时间的分布抽样，我们可以通过计算 $E(f(\tau))$ 来对相关性交易定价，其中 $\tau=(\tau_1, \cdots, \tau_k)$ 是违约时间的向量，f 是计算折现现金流的函数． ◀

在拟合时，由于各路径上债务人违约的时间是确切已知的，我们可以得到准确的估值，并能够准确地度量风险，无论多资产的支付结构多么复杂．许多研究都考虑了随机回收率以及回收率与违约的负相关关系，从而更好地对信用衍生工具估值．

和所有拟合方法一样，我们需要考虑计算量的问题，尤其是要进行灵敏度测量的时候. 已知违约通常属于不常见事件，那么重要性采样和低差异序列可以减小噪声和方差．

可以出于快速计算灵敏性度量的需要，产生了所谓的半解析方法．做到这一点需要完成几个步骤．其中一步是建立单因子模型，并认为违约只可能发生在有限个时间点．就精确性的需要而言，这样的假设就足够了．

4. 一个半解析单因子高斯 Copula 模型

这里我们利用一个单因子模型来建立风险中性损失分布，并介绍如何使用不同期限的损失分布序列来对合成损失份额定价．

假设第 j 个发行者(来自今天或未来某已知时间)用一个标准正态随机变量表示，其均值为 0，标准差为 1，满足如下形式：

$$a_j = \rho_j z_m + \sqrt{1-\rho_j^2} z_j \tag{23.68}$$

其中 $z_j(j=1, \cdots, d)$独立同分布，都服从 $N(0,1)$. z_m 是公共市场因子，受系统风险影响；z_j 是第 j 名发行者的具体风险或非系统风险 .

如果资产收益下降到某阈值 L_j 以下，那么表示风险事件发生 . 这意味着第 j 名发行者违约，此时 $a_j<L_j$. 已知 a_j 是标准正态变量，我们可以将违约事件写为

$$P_j = \Phi(L_j) \tag{23.69}$$

因此

$$L_j = \Phi^{-1}(P_i) \tag{23.70}$$

这意味着校正鞅概率只需通过对标准正态变量的分布函数求逆 .

392

在这种以 z_m 为条件的设定下，资产收益是相互独立的，因此我们能够很容易地计算条件违约概率 . 现在，在已知公共因子 $z_m=u$，u 是某个实数，我们可以看到第 j 个发行者将发生违约，只要

$$a_j = \rho_j z_m + \sqrt{1-\rho_j^2} z_j < L_j \tag{23.71}$$

已知 $z_m=u$，求解 z_j，可得

$$z_j \leqslant \frac{L_j - \rho_j u}{\sqrt{1-\rho_j^2}} \tag{23.72}$$

因此，第 j 名发行者的条件违约概率为

$$p_j(u) = \Phi\left(\frac{L_j - \rho_j u}{\sqrt{1-\rho_j^2}}\right) \tag{23.73}$$

如果假设每名发行者违约所造成的损失相同，将投资组合中的资产汇总，即可得到总损失的分布 . 可以根据以下步骤得到：

从单一资产开始，有两种可能，以 $p_1(u)$的概率损失 l 或者以 $1-p_1(u)$不损失 .

现在向投资组合加入第二项资产，调整损失 . 没有损失表示两种资产都没有违约，对应的概率为$(1-p_1(u))(1-p_2(u))$. 损失 l 表示资产一违约资产二存活，或者资产一存活资产二违约，对应的概率为 $p_1(u)(1-p_2(u))+p_2(u)(1-p_1(u))$，即 $p_1(u)+p_2(u)-2p_1(u)p_2(u)$. 损失 $2l$ 表示两项资产都违约，概率为 $p_1(u)p_2(u)$.

我们重复上面的步骤，不断添加资产，直到所有资产都加入到投资组合中 .

然后我们对 z_m 取不同的值，重复上面的过程 . 再对损失分布求积分，范围是所有 z_m 可能的值，损失分布显然是标准正态分布 .

我们可以看到，在这种方法中，没用用到拟合，因而也没有噪声的问题 . 我们还可以使用其他方法，包括快速傅里叶变换 .

违约时间模型和单因子模型都是高斯 Copula 模型，这意味着只要我们使用相同的相关矩阵和鞅，结果将会是相同的 .

23.3.2 相关性产品的估值

之前已经介绍了多种处理违约相关性的方法，我们可以在相关性产品定价中使用这些方法．我们下面将研究的相关性产品为篮筐式违约互换以及合成损失份额．

这些相关性产品的支付不能用单一资产集合的信用违约情况来复制．我们已经说过市场参与者在对具有相关性的产品估值时，一般采取标准无套利原则．因此在对这些工具估值时，需要考虑所有违约可能性，并计算风险中性期望．

1. 一篮子违约互换

一篮子违约互换实际上是一种特质产品．其含义是违约资产的特征是已知的．有很多方法可以对违约篮子定价．其中一种是违约时间模型．

我们用拟合产生违约时间．由于已知拟合路径，投资组合中各资产违约时间也是已知的，我们可以按照以下步骤计算溢价．

首先将违约时间按照增序排列，将排列后的第 k 个违约记为 $\tau_k(j)$，同时表示第 j 个资产发生了违约．这里 j 是资产的编号．

然后我们计算每条路径保费付款中一基点的现值，时间截止到违约时间 $\tau_k(j)$ 和到期日 T 的较小值． 393

完成上一步计算后，对所有路径求平均，得到篮子 PV01.

显然如果 $\tau_k(j)$ 大于 T，那么第 j 项资产在到期日前一直“存活”，这种情况所对应的现金流现值为 0. 否则现金流的现值是 $B(\tau_k(j))(1-R(j))$，其中 B 是 Libor 折现因子，$R(j)$ 是第 j 项资产违约时的回收率．根据这样的现金流计算方式，计算出各条路径上产生现金流的平均值．然后用篮子资产的 PV01 去除保护价值得到溢价的公平价格．

该方法的优点是便于实施且计算较快．Gregory 和 Laurent(2005)介绍了一种解析方法，能够在保持违约资产身份的情况下建立篮子中的第 n 次违约概率．和蒙特卡洛模拟相比，该方法能够使计算时间明显地下降．此外该方法参数的数量较少，便于校准．

2. 合成损失份额定价

根据拟合路径，我们已知投资组合内各资产违约的时间，为了对损失份额估值，我们将进行以下步骤：

对每条路径，我们计算本金损失的现值．

计算保费中 1 基点的现值，即 PV01.

求所有路径 PV01 的平均值，求保护价值的平均现值．用保护现值除以份额 PV01，得到所谓的平衡份额溢价．

该方法很简单，但如果投资组合中资产数量较多，那么计算量较大，速度慢．有一种半解析法可以成为该方法的替代方法．半解析法是可行的，因为标准合成损失份额可以通过损失分布求出．

23.4 期权市场中的信用溢价

研究本节内容的动机来自于企业和信用市场的内在关系．为了搞清楚这一点，我们可以考虑以下情况．某投资者认为某公司前景黯淡，很有可能违约．该投资者可以通过信用市场或者 CDS 合约、权益投资市场、看跌期权表现这种观点．如果想将这种观点以成本最低的方式应用到权益投资市场，我们可以建议该投资者购买深度虚值看跌期权．

如果违约发生，有两种可能性：预期违约；非预期违约．

在预期违约中，标的过程扩散到 0，就像反双曲正弦正态模型．在非预期违约中，标的过程跳跃到 0，正如 Merton 模型．

我们来考虑非预期违约．我们假设标的股票服从纯跳跃过程，在 0 附近有无限的活跃度．这种过程可以反映非预期违约，可以确保我们一定能够求出短期债券的溢价，这一点与扩散过程不同．

方差伽马(VG)属于这类过程．有许多此类列维过程，如具有随机到达的方差伽马过程(VGSA)、CGMY 等．

本节我们将介绍列维过程对于期权的基本校正，以及如何使用传统违约模型中的信息．我们将考虑以下模型：Merton 违约模型；风险率方法；Longstaff-Schwartz 模型．394

此外，我们将回顾 Hirsa 和 Madan 提出的一种简单方法，该方法能够估计期权市场的信用溢价．对于这 4 种情况，我们将介绍推导过程并给出一些例子．

23.4.1 修正的 Merton 违约模型

设 V 为公司价值，E 为权益价值，F 为债务面值．这里我们假设公司价值 V 的动态变化与权益价值 E 近似相同．到期日风险债券的价值，等于债务面值与公司价值中的较小值．也就是说

$$\text{到期日债券的价值} = \min(F,V) \tag{23.74}$$

$$= F - \max(F-V,0) \tag{23.75}$$

折现之后，即可得到风险/可违约债券的当前价值，我们可以看到可违约债券的价值为

$$\text{可违约债券的价值} = F\mathrm{e}^{-rT} - \text{看跌期权价格} \tag{23.76}$$

这意味着

$$F\mathrm{e}^{-(r+h)t} = F\mathrm{e}^{-rt} - \mathrm{Put}$$

$$F\mathrm{e}^{-rt}(1-\mathrm{e}^{-ht}) = \mathrm{Put}$$

$$\mathrm{e}^{-ht} = 1 - \frac{\mathrm{Put}}{F\mathrm{e}^{-rt}}$$

这里我们可以提出一些问题．如何从期权市场提取信息？如何计算债务面值和公司价值？

我们可以简单地假设公司价值服从纯跳跃过程，如方差伽马，并对看跌期权定价．

$$\mathrm{e}^{-ht} = 1 - \frac{\mathrm{Put}}{F\mathrm{e}^{-rt}}$$

$$= 1 - \Phi(-d_2) + \frac{Ve^{rt}}{F}\Phi(-d_1)$$
$$= 1 - \Phi(-d_2) + (\alpha + \beta s)\Phi(-d_1)$$
$$= f(\alpha,\beta,\sigma,\nu,\theta,t)$$

在校正过程中，目标是使下式最小化

$$\sum_{t=1}^{n} \mid e^{-ht} - f(\alpha,\beta,\sigma,\nu,\theta,t) \mid \tag{23.77}$$

从而得到 $\alpha,\beta,\sigma,\nu,\theta$.

如何计算看跌期权价值

$$\text{Put} = E(e^{-rT}\ (F - V_T)^+)$$

其中参数含义与之前相同，F 为债务面值，V 为公司价值. Merton(1973)证明，股票可以视为公司的看涨期权，执行价格等于债务面值．因此，在第 u 年，我们有

$$S_u = C(V_u) = E(e^{-r(T-u)}(V_T - F)^+) \tag{23.78}$$

通过反函数，我们可以写出

$$V_u = C^{-1}(S_u) \tag{23.79}$$

根据期权平价关系我们有

$$P_u = C(V_u) + Fe^{-r(T-u)} - V_u \tag{23.80}$$
$$= S_u + Fe^{-r(T-u)} - V_u \tag{23.81}$$
$$= S_u + Fe^{-r(T-u)} - C^{-1}(S_u) \tag{23.82}$$

395

假设 S_u 的分布已知，我们有

$$\text{Put} = E_0(P_u) \tag{23.83}$$
$$= \int_0^{\infty} e^{-ru} P_u f(S_u) dS_u \tag{23.84}$$
$$= \int_0^{\infty} e^{-ru}(S_u + Fe^{-r(T-u)} - C^{-1}(S_u)) f(S_u) dS_u \tag{23.85}$$

因此对于 Merton 违约模型，T 年的远期溢价为

$$h = -\frac{\ln\left(1 - \dfrac{\int_0^{\infty}(S_t + Fe^{-r(T-t)} - C^{-1}(S_t)) f(S_t) dS_t}{F}\right)}{T}$$

23.4.2 股权依赖风险(EDH)率方法

假设风险率[1]具有以下结构 $h(u) = a - b\ln S_u$，其中 S_u 服从纯跳跃过程．[2]可违约债券

---

[1] 该风险率方法是 Madan 和 Unal(2000)中双因子风险率模型的一般形式，其作用为对风险债务定价以及求信用溢价的期限结构．

[2] $S_u = S_0 e^{(r-q+\omega)u + x*u}$，其中 $x * u = \int_0^u \int_{-\infty}^{\infty} x\mu(dx, ds) du$.

的价格为

$$E\left\{\mathrm{e}^{-\int_0^t (r(u)+h(u)\mathrm{d}u)}\right\} \tag{23.86}$$

假设利率为常数，可得

$$\mathrm{e}^{-rt}E\left\{\mathrm{e}^{-\int_0^t h(u)\mathrm{d}u}\right\} \tag{23.87}$$

为找到它的封闭形式表达式，我们按照 Madan 和 Unal(2000)的方法，假设风险率 $h(u)$关于股票价格的对数是线性的，即 $h(u)=a-b\ln S_u$. 按照 Madan 和 Unal(2000)，这里我们假设 S_u 服从纯跳跃过程，我们可以写出 $S_u=S_0\mathrm{e}^{(r-q+\omega)u+x*u}$或者等价地

$$\ln S_u = \ln S_0 + (r-q+\omega)u + (x*\mu)_u \tag{23.88}$$

其中 $x*u=\int_0^u\int_{-\infty}^{\infty} x\mu(\mathrm{d}x,\mathrm{d}s)$,这里$\mu$是取值为正数的随机测度函数,统计$u$时刻幅度为$x$的跳跃的次数. 代入后从 0 到 t 对 $h(u)$ 积分可得

$$\int_0^t h(u)\mathrm{d}u = \int_0^t (a-b\ln(S_u))\mathrm{d}u \tag{23.89}$$

$$= \int_0^t (a-b(\ln S_0+(r-q+\omega)u+(x*\mu)_u)\mathrm{d}u \tag{23.90}$$

$$= (a-b\ln S_0)t-(r-q+\omega)b\frac{t^2}{2}-b\int_0^t (x*\mu)_u\mathrm{d}u \tag{23.91}$$

为计算(23.91)中的最后一项，我们首先改变积分的顺序，先对一部分积分，减少阶数，再使用 Fubini 定理，最后利用 * 的定义：

$$\int_0^t (x*u)_u\mathrm{d}u = \int_0^u\int_{-\infty}^{\infty} x\mu(\mathrm{d}x,\mathrm{d}s)\mathrm{d}u \tag{23.92}$$

$$= \int_{-\infty}^{\infty}\int_0^t\left(\int_s^t x\,\mathrm{d}u\right)\times\mu(\mathrm{d}x,\mathrm{d}s) \tag{23.93}$$

$$= \int_{-\infty}^{\infty}\int_0^t (t-s)\times x\mu(\mathrm{d}u,\mathrm{d}s) \tag{23.94}$$

$$= (t-s)x*\mu \tag{23.95}$$

将其带入(23.87)，求期望可得

$$E\left\{\mathrm{e}^{-\int_0^t h(u)\mathrm{d}u}\right\} = \mathrm{e}^{(a-b\ln S_0)t+(r-q+\omega)\frac{t^2}{2}}E[\exp(b(t-s)x*\mu)] \tag{23.96}$$

现在我们需要计算该期望. 我们需要使用跳跃过程的补偿量. 以下过程是鞅

396

$$(\mathrm{e}^{bx(t-s)}-1)*(\mu-\nu) \tag{23.97}$$

其中 ν 是鞅补偿量. 因此[⊖]

$$\mathcal{E}\left[(\mathrm{e}^{b(t-s)x}-1)*(\mu-\nu)\right] \tag{23.98}$$

⊖ 这里 ε 是随机指数. 更多细节参见 Jacod 和 Shirayev 的随机过程极限理论.

也是鞅．故

$$E\left\{\mathcal{E}\left[\left(e^{b(t-s)x}-1\right)*(\mu-\nu)\right]\right\}=1$$

或者等价地

$$E\left[e^{-\int_0^t\int_{-\infty}^{+\infty}\left(e^{b(t-s)x}-1\right)\nu(dx,ds)}e^{bx(t-s)*\mu}\right]=1$$

因此

$$E\left[e^{bx(t-s)*\mu}\right]=e^{\int_0^t\int_{-\infty}^{+\infty}\left(e^{b(t-s)x}-1\right)k(x)dxds}$$

其中 $k(x)$是标的过程的列维密度．已知方差伽马模型，我们有以下列维密度的表达式

$$k(y)=\frac{e^{-\lambda_p y}}{\nu y}\mathbf{1}_{y>0}+\frac{e^{-\lambda_n|y|}}{\nu|y|}\mathbf{1}_{y<0}$$

$$\lambda_p=\left(\frac{\theta^2}{\sigma^4}+\frac{2}{\sigma^2\nu}\right)^{\frac{1}{2}}-\frac{\theta}{\sigma^2}$$

$$\lambda_n=\left(\frac{\theta^2}{\sigma^4}+\frac{2}{\sigma^2\nu}\right)^{\frac{1}{2}}+\frac{\theta}{\sigma^2}$$

经过代数计算，可以得到以下期望表达式

$$e^{\int_0^t\int_{-\infty}^{+\infty}\left(e^{b(t-s)x}-1\right)k(x)dxds}=e^{\frac{2t}{\nu}}\left(1-\frac{b}{\lambda_p}t\right)^{-\frac{1}{\nu}\left(1-\frac{b}{\lambda_p}t\right)}\left(1+\frac{b}{\lambda_n}t\right)^{-\frac{1}{\nu}\left(1+\frac{b}{\lambda_n}t\right)} \tag{23.99}$$

$$\text{违约债券的价格}=e^{-rt}E\left\{e^{-\int_0^t h(u)du}\right\} \tag{23.100}$$

$$=e^{-\left(r+(a-b\ln S_0)-\frac{2}{\nu}\right)t+(r-q+\omega)\frac{t^2}{2}}$$

$$\times\left(1-\frac{b}{\lambda_p}t\right)^{-\frac{1}{\nu}\left(1-\frac{b}{\lambda_p}t\right)}\left(1+\frac{b}{\lambda_n}t\right)^{-\frac{1}{\nu}\left(1+\frac{b}{\lambda_n}t\right)} \tag{23.101}$$

我们观察到对于方差伽马模型，指数项的积分是明确的．对于期限内的每个时间点，我们都得到了 VG 参数(σ，ν，θ)通过校正和建立明确的信用溢价期限结构．我们可以展示出期权市场的偏度和峰度对违约溢价的影响．

23.4.3 Longstaff-Schwartz 模型

在该方法中，我们可以利用列维过程的首达时间来计算可违约权益(defaultable claim，参看 Longstaff 和 Schwartz(1995))．在寻找首达时间的过程中，我们首先定义

$$G(s,t,T)=E_t^Q\left[\mathbf{1}_{(s(u)<H,0\leqslant u\leqslant T)}\right] \tag{23.102}$$

该式是随机变量的条件期望，这里若公司的股票价格低于某事先设定的阈值 H，则表示会发生破产或债券的违约．从构造来看 $G(s,t,T)$是鞅，因为它是终端随机变量的条件期望．而且 $G(s,t,T)$是以下 PIDE 的解：

$$G_t+(r+\omega)sG_s+\int_{-\infty}^{\infty}(G(se^y,t,T)-G(s,t,T))k(y)dy=0$$

并满足以下边界条件：

$$G(s,t,T)=1 \quad 若\ s<H, \quad 对所有\ t$$

397

$$G(s,T,T)=0 \quad 若\ s \geqslant H$$

对于方差伽马、CGMY、逆高斯模型，我们有列维密度 $k(y)$，可以求解 PIDE 并得到参考实体的违约概率．

23.4.4 期权价格隐含的信用溢价——一个简单模型

首先我们假设价格比的对数 $z=\log(S_t/S_0)$服从随机过程[⊖] $h(z)$. 设 p 为以时间 t 为界限的违约概率，即参考实体在 t 之前违约的概率为 p. 那么 $1-p$ 为该参考实体的生存概率. 只要该公司在 t 之前不违约，价格比的对数的分布即为 $h(z)$. 若在 t 之前违约，则以概率 1 取值为 0. 定义

$$f(z;p)=\begin{cases}h(z) & 概率为\ 1-p \\ 1 & 概率为\ p\end{cases} \tag{23.103}$$

其中 p 为 t 时刻前违约概率．在该假设下，我们有另一种方法对看涨/看跌期权定价．看涨期权价格可用如下方法计算

$$\widetilde{C}(S_0,K,t)=\mathrm{e}^{-rt}\int(S_t-K)^+ f(z;p)\mathrm{d}z \tag{23.104}$$

$$=\mathrm{e}^{-rt}(1-p)\int(S_t-K)^+ h(z)\mathrm{d}z+p\int(0-K)^+ \mathrm{d}z \tag{23.105}$$

$$=(1-p)\mathrm{e}^{-rt}\times\int(S_t-K)^+ h(z)\mathrm{d}z \tag{23.106}$$

$$=(1-p)C(S_0,K,t) \tag{23.107}$$

看跌期权也可以用同样的方法计算：

$$\widetilde{P}(S_0,K,t)=\mathrm{e}^{-rt}\int(K-S_t)^+ f(z;p)\mathrm{d}z \tag{23.108}$$

$$=\mathrm{e}^{-rt}\left((1-p)\int(K-S_t)^+ h(z)\mathrm{d}z+p\int(K-0)^+ \mathrm{d}z\right) \tag{23.109}$$

$$=(1-p)\mathrm{e}^{-rt}\int(S_t-K)^+ h(z)\mathrm{d}z+pK\mathrm{e}^{-rt} \tag{23.110}$$

$$=(1-p)P(S_0,K,t)+pK\mathrm{e}^{-rt} \tag{23.111}$$

利用方程(23.107)或(23.111)，根据(看涨或看跌期权)市场价格校正后，我们可以得到概率 p. 计算出 p 之后，我们就可以计算信用溢价 s.

通过这种方法，我们可以通过期权价格直接得到信用溢价．在方差伽马模型中，我们称 VG 过程*跳跃至破产*．对于每个到期日 t，我们得到 p，并利用 FFT 校正波动率平面，得到方差伽马模型参数 σ、ν 和 θ.

为什么选择方差伽马过程？

以上模型的一个重要部分是我们可以自由选择标的过程 $h(z)$，也就是违约出现时资产服

⊖ 这包括方差伽马(VG)、CIR、CEV、Merton 跳跃扩散、Kou 跳跃扩散模型等．

从的过程．方差伽马是一个很好的选择，因为如果使用方差伽马，我们不仅可以校正均值和方差，还可以校正峰度和偏度．我们非常关注违约概率的信息，因而在选择分布时，能否准确地求出分布的高阶矩非常关键，高阶矩可能包含市场中隐含违约风险的重要信息．

23.4.5 小结

本节为了从期权市场提取信息来估计违约概率，我们介绍了以下方法： 398

- 利用列维过程的首达时间，计算 Longstaff 和 Schwartz 所定义的可违约权益．
- 列维过程对期权面的基础校正，以及 3 个传统违约模型的信息的使用（Merton，Longstaff-Schwartz，风险率）．

23.5 习题

1. 继续风险率方法对违约债券价格的推导，找到可违约债券的表达式．假设标的过程服从 CGMY 过程．CGMY 的列维密度为

$$k(y)=\frac{\mathrm{e}^{-\lambda_p y}}{\nu y^{1+Y}}\mathbf{1}_{y>0}+\frac{\mathrm{e}^{-\lambda_n|y|}}{\nu|y|^{1+Y}}\mathbf{1}_{y<0}$$

$$\lambda_p=\left(\frac{\theta^2}{\sigma^4}+\frac{2}{\sigma^2\nu}\right)^{\frac{1}{2}}-\frac{\theta}{\sigma^2}$$

$$\lambda_n=\left(\frac{\theta^2}{\sigma^4}+\frac{2}{\sigma^2\nu}\right)^{\frac{1}{2}}+\frac{\theta}{\sigma^2}$$

2. 叙述 Bootstrap 过程．
3. 在例 7 中，找到平衡溢价，使得固定保费的现值等于或有性补偿的现值．
4. 利用 23.4.4 节所介绍的 Hirsa-Madan 模型，校正下表中沃尔玛的虚值看跌期权的保费．[⊖] 期权市场对该公司的违约概率看法如何？假设沃尔玛的回收率是 40%，使用信用三角公式，估计信用溢价．假设你已知市场上信用溢价是 50 基点，评价期权所隐含的信用溢价的准确度，以及股票期望和信用市场的一致性．

K	bid	ask	t	r
30	0.04	0.07	0.461 75	0.0001
32.5	0.06	0.09	0.461 75	0.0001
35	0.08	0.12	0.461 75	0.0001
37.5	0.11	0.14	0.461 75	0.0001
40	0.15	0.19	0.461 75	0.0001
42.5	0.21	0.24	0.461 75	0.0001
45	0.28	0.32	0.461 75	0.0001
47.5	0.41	0.44	0.461 75	0.0001
50	0.61	0.63	0.461 75	0.0001

⊖ 数据来自 OptionMetrics.

（续）

K	bid	ask	t	r
52.5	0.91	0.94	0.461 75	0.0001
55	1.38	1.42	0.461 75	0.0001
57.5	2.1	2.13	0.461 75	0.0001

5. 考虑首次违约篮子(FTD)，对篮子内的首次违约支付$(1-R)$，篮子内含 3 项参考信用．该 FTD 的保护期限为 5 年，保费按季度支付．合约规定年溢价为 800 基点．3 种参考信用的常数风险率分别为 4%、6%、10%．每种信用的回收率为 0.5，相互之间的相关系数为$\rho=0.2$.使用高斯 Copula 模型，模拟违约时间，计算合约现值．假设利率为常数 3%.

399
取 $\rho=0.5$，重复上述计算．评价违约相关系数与 FTD 价值的关系．

第 24 章　停时与美式证券

24.1　引言

本书所介绍的期权可分为两类．从定价方程来看，第一类的特点是期权价值只与标的资产现值 S_t 和时间 t 有关．例如，t 时刻普通看涨期权的价值可写为

$$C(t) = F(S_t, t) \tag{24.1}$$

已知 t 时刻以及观察值 S_t，期权价格可由以上函数确定．普通欧式期权就属于这一类，其中 S_t 是一个几何过程．㊀

第二类期权具有路径依赖性，在本书中没有具体介绍．这类期权在 t 时刻的价格不仅依赖于当前的 S_t，可能还依赖于过去某些或过去所有 S_t 的值．例如，在到期日 T，期权的支付额可能取决于离散观察点上观察值的平均：

$$t < t_1 < t_2 < \cdots < t_N = T \tag{24.2}$$

在到期日，看涨期权持有者可能支付：

$$C_T = \max\left[\frac{S_{t_1} + S_{t_2} + \cdots + S_{t_N}}{N} - K, 0\right] \tag{24.3}$$

其中 K 是某个执行价格．

在这些条件下，T 时刻该看涨期权的价格可以写为：

$$C_T = F(S_{t_1}, S_{t_2}, \cdots, S_{t_N}) \tag{24.4}$$

显然，在 t 时刻看($t<T$)，该表达式可能十分复杂．㊁然而，对这种奇异期权定价不一定比普通奇异期权更困难．实际上，根据我们前面的内容，该期权的支付在到期日进行，从这种角度来说它仍是欧式期权．唯一复杂的地方是表达式中增加了 S_{t_i} 项．因此，尽管这种期权是路径依赖的，其付款依赖于标的资产价格与到期日的价格之间的路径．只要确定了 S_t 的变化规律，蒙特卡洛方法就可以有效地估计 C_t.

注意到对以上两种期权，投资者在购买后都无需对期权的执行再做抉择．在这两种情况下，投资者都要等到到期日才行使权利(如果期权实值的话)．或者，期权持有者可以平仓并将期权卖给他人．但是此外再没有别的选择．因此，公式中没有加入别的变量．㊂

现在考虑美式期权．这种证券可以在到期日 T 或者提前被执行．一旦投资者购买了美式期权，他将面临一个新的问题．他必须选择一个合适的时间来执行期权，而不能仅仅坐等到期日的到来．在某个记作 θ 的时刻，$\theta \in [0, T]$，行使该期权可能使他获利更多，也就

㊀ 其他假设还包括无红利和利率为常数，从而得到 $F(\cdot)$ 封闭形式的解．

㊁ 实际上，不可能存在封闭形式的解．

㊂ 我们总假设利率和波动率为常数．

是说

$$S_\theta - K \tag{24.5}$$

可能要高于到期日的期权价值

$$\max[S_T - K, 0] \tag{24.6}$$

事实上在某特定时刻 θ，在鞅测度下，未来收益 $\max[S_t - K, 0]$的期望可能小于投资者立即执行期权所得到的收益 $S_\theta - K$. 也就是说，在即期利率为常数的条件下，我们可能有：

$$S_\theta - K > E^Q[\mathrm{e}^{(T-\theta)r}\max[S_T - K, 0] \mid I_\theta] \tag{24.7}$$

该式表示期望收益的折现值可能小于投资者在 θ 时刻行使期权得到的收益．

从这一点可以看出，研究何时行使美式期权等价于寻找这种关键时刻 θ. 注意到在这些条件下，该期权的定价公式可能依赖于选择 θ 的方法，以及列入考虑范围的变量．

这种 θ 叫做停时．当用某种优化方法确定了执行日期后，所确定的时间叫做终止时间，它们在美式证券定价中起着关键的作用．

24.2 为什么研究停时

402

虽然停时的概念局限于美式证券，可是我们很有必要学习这一内容．确实，大多数金融衍生品都是美式的，停时是必不可少的定价工具．然而它还有更多的意义．我们学习停时不仅是因为它是理论公式中使用的概念，还因为它是提前结束日期的数值算法中重要的概念．也就是说，我们学习停时也是基于数值计算方面的考虑．

最优停时一些性质使某些方法在定价时更为方便．学习这些性质后，我们可以减少计算时间，即决定在某时刻 t^* 要不要执行某期权．或者说，对 θ 而言，是否有

$$\theta = t^* \tag{24.8}$$

该式表示执行．或者

$$\theta > t \tag{24.9}$$

该式表示不执行．在此类计算更快捷准确之后，投资者可以降低成本，更有效地捕捉套利机会．因此，选取合适的算法在定价中起着重要作用．

此外研究停时还有一些原因．最优停时通常是使用所谓动态规划法得到的．动态规划本身是一项非常有用的工具，无论是否研究衍生品定价，都应对其有所了解．很巧的是，最优停时的求解非常适合动态规划法．

美式证券

美式衍生证券包括隐含或显式期权，可以在到期日前提前执行．这不但使求证券公平市场价值更复杂，也使得期权价格的计算更复杂．

百慕大期权是美式与欧式期权的结合．它可以在到期日前某些提前约定的时间执行．但是它并不能在$[0, T]$之间的任意时刻执行．在发行时，合约就会具体规定一些日期 $t_1 < t_2 < \cdots < t_n = T$，期权持有者可以在这些时刻执行期权．

从最优停止时间的角度看，百慕大期权的复杂程度和美式证券相似．我们对停时和相关工具所介绍的内容如果对美式期权有效，对百慕大期权也同样有效．因此，本章我们只研究美式期权的停时．

24.3 停时

停止时间是一种特殊的随机变量，表示一个随机的时刻．例如，设 τ 为停时．那么这意味着两方面内容．第一，τ 是随机的．第二，τ 的取值范围是 $[0,T]$，$T>0$ 为已知．当结果出现时，可表示为

$$\tau = t \tag{24.10}$$

也就是说，该随机变量的结果是一个特定时刻．

现在考虑建立在债券上的美式期权．该期权可以在当前时刻 $t=0$ 与到期日 T 之间的任何时刻执行．如果期权执行者认为时机已到，那么他将执行期权，否则他将选择等待直到到期．

所以，我们需要研究一个“随机日期”，它对于期权定价非常重要．实际上，提前执行的权力为美式期权增加了额外的价值． 403

我们设 τ 为期权的执行日期．很显然已知信息集 I_t，我们就可以知道该期权是否已经被执行．也就是说，已知 I_t，我们可以知道哪种情况成立：

$$\tau \leqslant t \tag{24.11}$$

该式表示期权已经被执行，或者

$$\tau > t \tag{24.12}$$

该式表示期权尚未被提前执行．

τ 的这种性质可以用来定义停止时间．

定义 27 停止时间是 I_t 可测的非负随机变量，满足

1)已知 I_t，我们可以判断是否有

$$\tau \leqslant t \tag{24.13}$$

2)我们有

$$\Pr(\tau < \infty) = 1 \tag{24.14}$$

对于一般的衍生证券，期限都是有限的．因此期权要么在一个有限的时刻被执行，要么过期失效．这意味着第二个等式，即 τ 以 1 概率有限，将恒成立．

24.4 停时的作用

在实际操作中，我们应该如何利用停时 τ 呢？

最明显的一个用途是用 τ 来代表期权的执行时间．对欧式期权而言，执行时间不含随机因素，期权只能在到期日被执行．因而我们有

$$\Pr(\tau = T) = 1 \tag{24.15}$$

而对于美式期权，τ 一般是随机的．[1]

考虑一份美式看涨期权 $F(S_t, t)$，标的资产为 S_t，S_t 满足以下 SDE：

$$dS_t = a(S_t,t)dt + \sigma(S_t,t)dW_t, \quad t \in [0,\infty) \tag{24.16}$$

其中漂移项和扩散项系数满足一般的规范性条件．

和过去一样，衍生证券的价格可以在等价鞅测度 Q 下表示．但此时多了一个复杂之处．证券持有者不一定要等到 T 时刻才执行期权，他可以在利润较高时提前执行．

也就是说，如果持有者必须等到到期日才能执行期权，那么该期权 t 时刻价值为

$$F(S_t,t) = E_t^Q[e^{-r(T-t)}\max\{S_T - K,0\}] \tag{24.17}$$

而如果他可以提前执行，那么价值为

$$F(S_t,t^*) = \sup_{\tau\in\Phi_{t,T}}[E_\tau^Q[e^{-r(T-\tau)}F(S_\tau,t,\tau)]] \tag{24.18}$$

其中 $\Phi_{t,T}$ 是所有可能停止机会的集合，[2] t^* 是 τ 的最优选择．这里 τ 代表期权持有者可能执行该看涨期权的日期．

因此在 t 时刻，我们可以计算出一系列 $F(S_\tau,t,\tau)$，包含所有停止时间 τ 可能的取值．为了求出正确的价格，我们要求这些 $F(S_\tau,t,\tau)$ 的上确界．

404

24.5 简化的设定

下面我们用一个简化的二叉树模型来研究美式看涨期权定价，以及最优停时．尽管这个设定很“简单”，而且我们此时的主要目标是理解与停时相关的工具，但在美式期权的实际定价中，所采用的方法也和我们此时所考虑的模型类似．因此，以下讨论从一些简单数值定价计算的角度来看，也是非常有用的．

模型

该二项模型假设标的资产 S_t 在连续时间下服从以下集合维纳过程：

$$dS_t = (r-\delta)S_t dt + \sigma S_t dW_t, \quad t \in [0,\infty) \tag{24.19}$$

其中 r 是常数瞬时即期利率，$\delta>0$ 是已知的红利率，W_t 是在风险中性测度 Q 下的维纳过程．

我们设 C_t 为美式看涨期权价格，K 为执行价格，T 为到期日，标的资产价格为 S_t．下面我们将用二叉树方法对该看涨期权定价．

该方法我们之前已经介绍过，方便起见，我们再总结一下．首先我们选择间距参数 Δ，对 S_t 用标准方法离散：[3]

$$S_i^u = S_{i-1}e^{\sigma\sqrt{\Delta}} \tag{24.22}$$

---

[1] 在一些特殊情况下，美式期权在期限内一直不值得执行，那么相应的 τ 还是会等于 T.

[2] 也就是说，它是 τ 所有可能结果的集合．

[3] 这是一种可行的离散方法．还有别的方法，比如我们可以令：

$$S_i^u = S_{i-1}e^{\left((r-\delta)-\frac{1}{2}\sigma^2\right)\Delta+\sigma\sqrt{\Delta}} \tag{24.20}$$

$$S_i^d = S_{i-1}e^{\left((r-\delta)-\frac{1}{2}\sigma^2\right)\Delta-\sigma\sqrt{\Delta}} \tag{24.21}$$

$$S_i^{\mathrm{d}} = S_{i-1}\mathrm{e}^{-\sigma\sqrt{\Delta}} \tag{24.23}$$

这里假设上移和下移的概率为常数，n 步中各状态之间都是如此，概率为：

$$\Pr(u) = \frac{1}{2} + \frac{(r-\delta)-\frac{1}{2}\sigma^2}{2\sigma}\sqrt{\Delta} \tag{24.24}$$

$$\Pr(\mathrm{d}) = 1-\Pr(u) \tag{24.25}$$

也就是说，一旦过程到达点 S_{i-1}，那么下一点要么是 S_i^u 要么是 S_i^{d}，相应的概率为 $\Pr(u)$ 和 $\Pr(d)$. ⊖

通过对离散参数的调整，随着 $\Delta\to 0$，该离散系统收敛到几何过程．也就是说，漂移和扩散参数将保持不变，而 S_t 的轨迹也还是按照(24.19)中的 SDE 变化．

注意上移和下移的概率参数 u，d 为常数，且可表示为

$$u = \mathrm{e}^{\sigma\sqrt{\Delta}} \tag{24.27}$$

$$d = \mathrm{e}^{-\sigma\sqrt{\Delta}} \tag{24.28}$$

此外，和往常一样，

$$ud = 1 \tag{24.29}$$

因此该二叉树在节点上是重合的．

图 24-1 展示了 S_t 的可能路径．我们来观察树的结构．水平方向的运动表示 S_i 随时间的变化．该过程从 S_0 出发，结束于 6 个到期状态中的一个．而当 $i=1,2,\cdots,5$ 时，S_i 可沿多种路径移动．实际上，总共有 2^n 种可能的路径，n 代表阶段的数量，$n=T/n$. 在这个例子中，S_i 共有 32 种可能的路径． 405

看涨期权的价格取决于 S_i 的路径．欧式看涨期权的情况我们之前已经讨论过．对于美式看涨期权，我们要采取完全不同的方法来观察路径，而看涨期权价格 C_n 对 S_i 路径的依赖关系也完全不同．

二叉树的标准形式如图 24-1 所示，水平方向的路径模拟了 S_i 随时间变化的行为．如果要分析停时、理解实时决定停止还是继续的复杂性，我们有必要换一个角度来观察该二叉树，即将二叉树旋转 90 度．这一开始可能会带来不便，但它能够极大促进我们对停止时间的相关数学工具的理解．

我们不再观察图 24-1 中的二叉树，转而考虑图 24-2. 横轴上我们标出了 S_i 从 $n=0$ 到 $n=5$所有可能的取值．在这段时间内，S_i 有 11 个可能的取值．记所有取值的集合为 E，利用条件 $ud=1$，我们有 406

$$E = \{u^5S_0, u^4S_0, u^3S_0, u^2S_0, u^1S_0, S_0, d^1S_0, d^2S_0, d^3S_0, d^4S_0, d^5S_0\} \tag{24.30}$$

现在，尽管从时间的角度看该二叉树和图 24-1 的相同，但当我们把它看作 S_i 可能取值的函数时，我们可以看到一些不一样的内容．这些值标在了图 24-2 的横轴上，这些点恰是集

⊖ 随着 $\Delta\to 0$，上移和下移的概率将等于

$$\Pr(u) = \Pr(d) = \frac{1}{2} \tag{24.26}$$

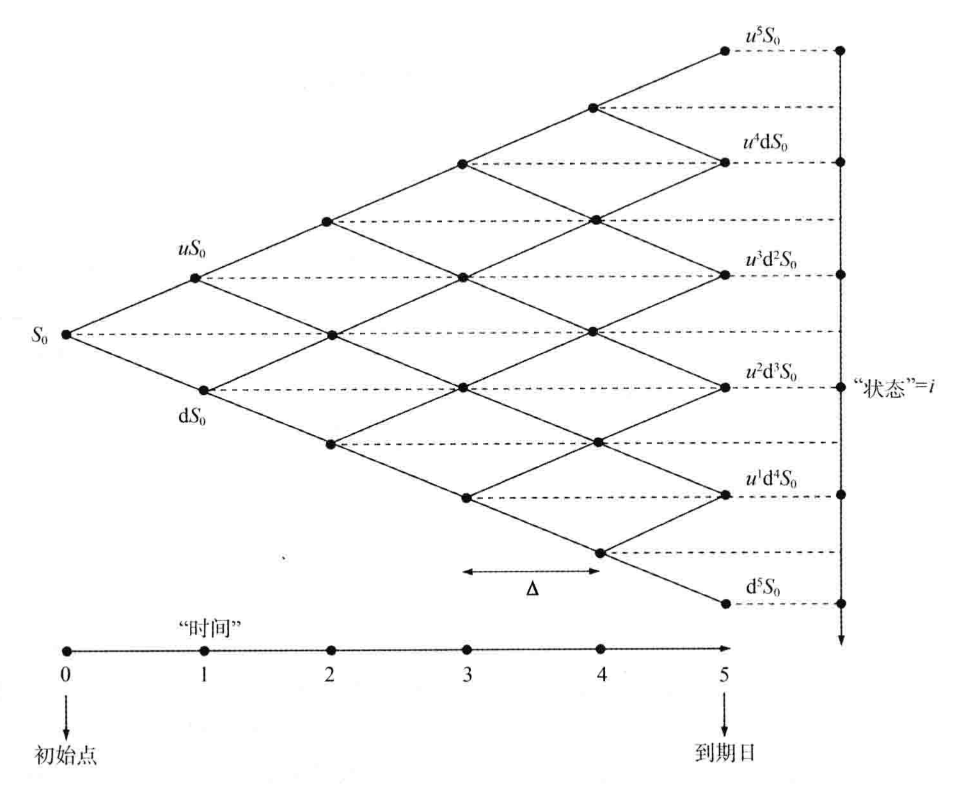


图 24-1 二叉树水平模拟股票随时间变化行为的标准方法

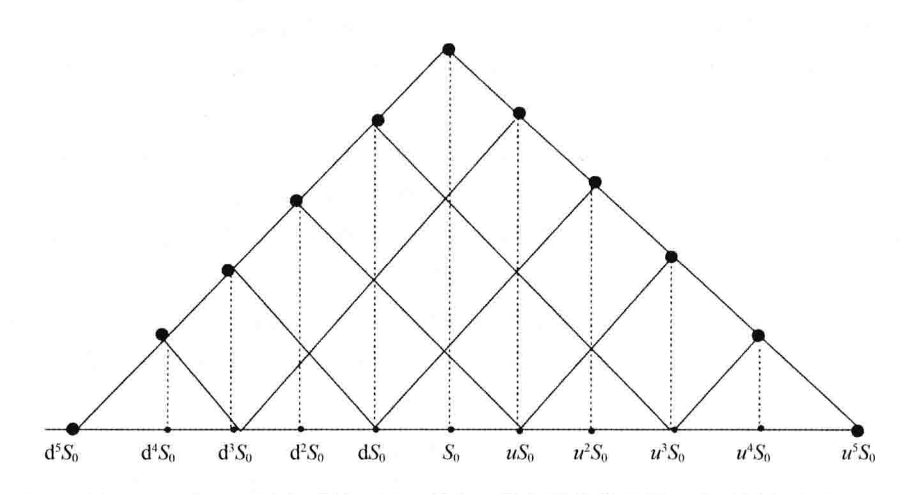


图 24-2 在二叉树水平轴上标记所有可能股票价值的另一种不同的方法

合 E 的所有元素．图 24-2 和图 24-1 完全相同，只是旋转了 90 度．

在美式期权中，我们可以提前执行期权．是否提前执行取决于 S_i 落在什么点．现在考虑过程在集合 E 上的表现，如图 24-2 所示．

一开始，我们位于中间点 S_0. 当下一时刻 $n=1$ 时，S_1 有 0.5 的概率左移到 dS_0，有 0.5 的概率右移到 uS_0. 移动之后，下一时刻 S_2 有 0.5 的概率回到 S_0 点，有 0.5 的概率继续向左或向右移动. 因此，在每一点，该过程只能移动到相邻的状态. 只有两个端点是例外，一旦过程到达该位置，则必须停止，因为需要花 5 个时间段到达端点，这恰好等于期权的期限.[㊀] 这意味着 S_i 还可以表示马尔可夫链在阶段 i 的位置.

现在记住，我们可以在任一阶段停止试验，只要我们希望这么做. 如果停止，我们的收益为:

$$S_i - K \tag{24.31}$$

如果我们决定继续实验，那么我们将持有价格为 $C(S_i)$的期权.

现在我们考虑如何做出最优决策. 设 τ 为我们提前执行期权的时间. 在初始时刻 $i=0$ 时 τ 是随机的，因为我们是否停止取决于 S_i 的轨迹.

407

假设我们考虑在阶段 i 停止，那么，令

$$\tau = i \tag{24.32}$$

显然做出该决策前，我们观察了 S_i 在阶段 i 之前的路径. 也就是说，我们有以下观察值:

$$\{S_0, S_1, \cdots, S_i\} \tag{24.33}$$

而停止决策是这些历史信息的函数. 根据这一点可知，提前执行期权的决策不依赖于对阶段 i 之后的状态的了解. i 以后的“未来”是未知的. 这就是“τ 是 I_i 可测”的含义.

现在，如果我们能够确定一种选择 τ 的策略，那么我们就能得到与随机变量 τ 相关的概率. 但如果无法确定这种策略，就无法了解 τ 的性质，τ 也就不是定义明确的随机变量. 因此首要任务是确定选择 τ 的策略. 那么要做什么呢?

考虑以下规则，其中我们用展开的形式来写条件期望算子:

$$F(S_0) = \max_{\tau} E^Q[C(s_\tau) \mid S_0] = \max_{\tau} E^Q[\mathrm{e}^{-rT}\max[S_\tau - K, 0] \mid S_0] \tag{24.34}$$

根据这一点，当我们选择 τ，我们就是在最佳时刻停止，期权的价值达到最大.[㊁]

然后我们需要做两件事情. 首先，我们需要得到以下变量:

$$\max_{\tau} E^Q[C(S_t)] \tag{24.35}$$

其次，我们需要找到规则确定最优停止时间，使得

$$E^Q[C(S_t)] \leqslant C(S_{\tau^*}), \quad 对\ t > \tau^* \tag{24.36}$$

完成这一步后，最优决策为以下形式:

$$\tau^* = \min[k: S_k > B(k, S_k)] \tag{24.37}$$

其中 $B(k,S_k)$是最优执行边界，依赖于 k 和 S_k 的当前值(可能还包括过去值). 该边界是

㊀ 在马尔可夫链的相关术语中，称两个端点为吸收状态. 也就是说，一旦到达改点，则将以概率 1 停留.

㊁ 根据该设定，$F(\cdot)$是目标函数. 如果假设 $F(\cdot)$有界，随后步骤的所有技术性条件都将满足. 在实际的定价算法中，通常默认有界性成立.

待定的.⊖

24.6 一个简单的例子

为了理解某些关于停时问题更深层次的内容，我们来讨论一个简单但很重要的例子．回顾美式看涨期权持有者所面临的问题．在期权期限内的任何一个时刻，持有者有权提前执行期权．因此，在所有 $t\in[0,T]$，持有者要决定是否执行期权．但是这种决策比它表面看上去还要复杂得多．

在做出选择前，投资者必须考虑他在未来是否会提前执行期权．也就是说，在当天做出决定前，投资者还要估计他未来做出同样决策的概率．只有分析了期权可能的未来收益之后，才能决定是否继续．而未来提前执行期权的概率又关系到未来的未来是否提前执行，以此类推．最后，美式期权的持有者所面对的是牵涉所有时间段的复杂问题．

408

该如何做这种决策呢？有没有一些机械化的规则能帮助决策者决定是否提前执行呢？最后，我们该如何认识这些相互关联的复杂决策？

下面这个简单的例子将为这些问题提供线索．读者会注意到，本例所采用的方法与上一节讨论的二叉树模型相似．实际上我们使用同样的记号．

假设我们观察到连续的随机变量值．我们令 $n=5$，则 S_i 有 11 个可能的取值．这些值由以下按序排列的集合 E 给出：

$$E\{a_1,a_2,a_3,a_4,a_5,a_6,a_7,a_8,a_9,a_{10}\} \tag{24.38}$$

初始值 S_0 已知为 a. 该过程从 $i=1$ 开始观察，是马尔可夫的，且按照以下假设变化：

1）在阶段 i 假设过程 S_i 等于 E 中的某个值之后，下一阶段它只能向左移一步或向右移一步．不存在其余可能性．例如，在 $i=3$ 时我们有 $S_3=a_6$，那么 S_4 等于 a_5 或者 a_7.

2）其次，a_1 和 a_{10} 具有吸收性．如果过程到达这些位置，那么将以概率 1 停留．这些位置只可能在到期日达到．

下面介绍模型中相关的转移概率．

根据上面的介绍，转移概率为：

$$\Pr(S_{i+1}=a_{j+1}\mid S_i=a_j)=\frac{1}{2} \tag{24.39}$$

$$\Pr(S_{i+1}=a_{j-1}\mid S_i=a_j)=\frac{1}{2} \tag{24.40}$$

$$\Pr(S_{i+1}=a_1\mid S_i=a_1)=\Pr(S_{i+1}=a_{10}\mid S_i=a_{10})=1 \tag{24.41}$$

其他转移的概率为 0.

$$\Pr(S_{i+1}=a_m\mid S_i=a_j)=0,\quad |m|>j+1 \tag{24.42}$$

这意味着过程 S_i 只能移动到相邻的位置，不能跳跃．对于初始点，我们有

⊖ 我们来看如何解释期望值 $E^Q[S_\tau]$. 此处随机变量为 τ. 因此需要将 S_τ 的可能取值乘以 $\tau=k$ 的概率．但是还需要考虑其他东西，因为即使固定 τ，S_τ 仍是随机的．因而，在诸如 $E^Q[S_\tau]$的期望中，我们会用 S_t 的可能取值乘以 S_t 取这些值的概率．

$$\Pr(S_1 = a_5 \mid S_0 = a) = \frac{1}{2} \tag{24.43}$$

$$\Pr(S_1 = a_6 \mid S_i = a) = \frac{1}{2} \tag{24.44}$$

该情况如图 24-3 所示．横轴表示集合 E. 横轴上的箭头表示可能的移动方向，并标明了相应概率．注意如果该过程到达两个端点，那么停止(被吸收).

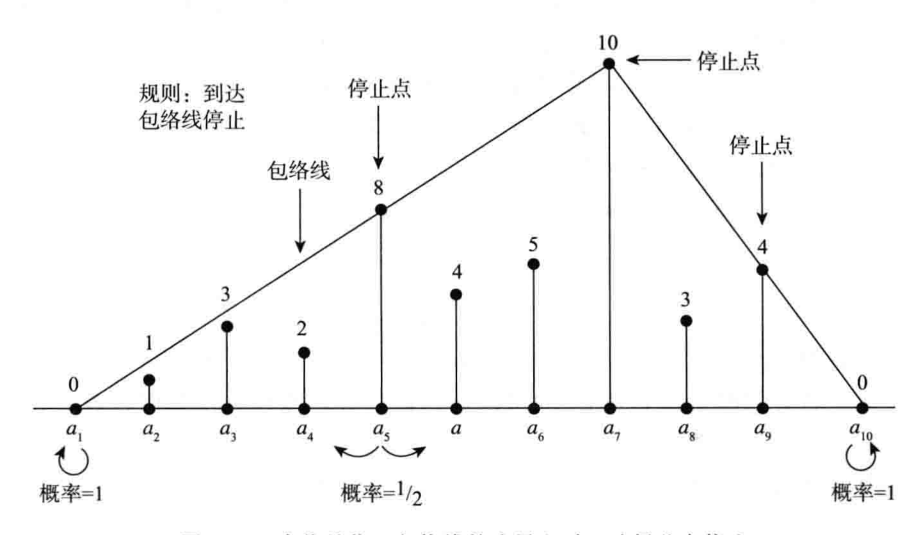


图 24-3　当收益位于包络线的边界上时，选择此点停止

为讨论最优停止决策，我们还需要介绍一点内容．当 S_i 到达某一位置，如 E 中的 a_j，投资者可以选择执行期权并得到收益 $F(a_j)$. 若投资者接受该收益，则游戏结束．若不接受，则游戏继续，S_i 将移动到相邻的位置．在图 24-3 中，每个 $a_j \in E$ 相应的收益用垂线表示在对应的点上．

决策者所面临的问题如下．他可以陆续观察到 S_i 并计算出相应的 $F(a_j)$. 决策者需要计算立即停止所得的收益，并与暂不停止的期望收益相比较．他应该怎么做呢？我们用图 24-3 来讨论最优决策．

注意，在每个位置，我们只知道当前我们在哪里．换句话说，我们能观察到当前值 S_i，但不知道未来的结果，尽管我们知道相应的概率．考虑这些可能．

409

假设我们到达了两个端点中的一个．显然我们没有选择，必须停止．该位置具有吸收性. 我们不会再移动到其他位置．

下面考虑位置 a_8. 我们是否应该在 $S_i = a_8$ 时停止，并获得收益 $F(a_8)$呢？答案是显然不应该(除非此时 $n=5$，我们必须停下). 下一次 S_i 要么移动到 a_7 要么移动到 a_9. 我们从图 24-3 可以看到，这两点的收益都比 $F(a_8)$要高．如果继续，我们肯定能够获得更高的收益，因而不能停止．

另一个明显的决策位于位置 a_7. 该位置具有最高的收益．一旦到达该点显然应该立即

停止．若选择继续，不会得到更好的结果．

这样看来，关于是否停止的决策一点也不复杂．再考虑两个位置 a_3 和 a_5. 此时决策就变得困难多了．这两点都具有以下性质．其收益是局部最高的．如果到达这些点之后选择继续，我们将到达相邻的位置，对应的收益更低．至少在即近的未来，持有者的处境会更糟．但如果不停止，他还保留获得更高收益的可能性，如 $F(a_7)$或 $F(a_5)$. 所以哪种决策更好？应该停止并获得局部最高 $F(a_3)$或 $F(a_5)$，还是继续持有，等待未来出现更高收益？

答案并不是一眼就能看出来的，我们需要自己评估未来可能性．实际上，a_3 和 a_5 两个位置答案是不一样的．应该在 a_3 选择继续，而在 a_5 选择停止．

从 a_5 开始介绍．假设在某个 i 我们有 $S_i=a_5$，如果选择停止，我们有 $F(a_5)=8$. 如果选择继续，我们的期望收益是多少呢？

很容易计算下一时刻的期望收益：

$$E[F(S_{i+1}) \mid S_i = 5] = \frac{1}{2}F(a_4) + \frac{1}{2}F(a) \tag{24.45}$$

410

$$= 3 \tag{24.46}$$

这明显低于我们当前能够锁定的收益 8. 但还有一点．游戏并不在下一阶段立即终止，如果只看下一阶段，就会忽略价格到达 a_7 的可能性，而这种可能性会使期望收益上升．

我们希望得到最优收益函数 $V(a_5)$，它能够等于两种收益中的较大值，即停止后的当前收益，或决定继续后的期望收益(假定我们用一种最优的方式继续). 也就是说，我们希望

$$V(a_5) = \max[E[\text{收益} \mid \text{停止}], E[\text{收益} \mid \text{继续}]] \tag{24.47}$$

这里选择停止的期望收益是已知的，就是 $F(a_5)$. 而选择继续的期望收益是未知的．若要计算，则需要对未来的状态用同样的方法考虑．

我们可以写出：

$$V(a_5) = \max[F(a_5), \frac{1}{2}V(a_4) + \frac{1}{2}V(a)] \tag{24.48}$$

注意这里假设不存在折现．

所以，在我们计算 $V(a_5)$之前我们要先计算 $V(a_4)$和 $V(a)$. 但是，相同的问题依然存在. 对应未来状态的 V 似乎仍是未知的．

这个推理看起来会陷入循环，实际上不会．在这个问题的设定中，有一些点的 $V(a_i)$ 是可以立刻求出来的．例如，我们刚才说过，当到达端点 a_1 或 a_7 时，有

$$V(a_7) = F(a_7) = 8 \tag{24.49}$$

以及

$$V(a_1) = F(a_1) = 0 \tag{24.50}$$

将它们带入(24.48)的“未来值”，我们最终可以求出 $V(a_5)$.

那么我们应该怎样利用此条件求(24.48)呢？我们可以写出

$$E[F(S_{i+1}) \mid S_i = a_7, \text{最优时刻停止}] = \frac{1}{2}V(a_4) + \frac{1}{2}V(a) \tag{24.51}$$

我们不再像(24.46)那样，只将下一个阶段的未来状态考虑进来，而是用 $V(a_4)$ 和 $V(a)$ 代替了(24.46)中的 $F(a_4)$ 和 $F(a)$. 这是因为后者没有考虑价格在以后达到高位并执行期权的可能性，而 $V(a_4)$ 和 $V(a)$ 则考虑了这一点．也就是说，$V(a_4)$ 和 $V(a)$ 包含了在未来进行最优化决策的思想．这种思想与衍生品定价时根据二叉树“倒推”的思路很类似．如果采取这种方法，很容易就能得出应该在 a_3 继续、a_5 停止的结论．注意，(24.51)左端的期望值与(24.46)不同，因为此时条件之一是我们在最优的时刻停止．这样得到的是真正的 $V(a_5)$.

最后关于这个例子说一点．图 24-3 中有一个有趣的现象．我们认为应该选择停止的点，其收益线段的上端点刚好位于我们标记为包络线的边界上．如果市场参与者已知该包络线，那么他选择最优停止时间的工作就会被大大简化．他所要做的只有观察收益是否落在包络线上．若当前收益等于包络线上该点的值，则停止，否则继续．本质上这就是方程(24.37)所表达的含义，后者给出了最优停止规则．

24.7 停时和鞅

在本章的最后，我们来了解一下鞅在停时的相关理论中所扮演的角色．大多数本书中的结果都可以拓展到停时方面的内容．下面我们简单给出两个这类结果，但不作进一步介绍． 411

24.7.1 鞅

假设 M_t 为在概率 P 下的连续时间鞅，有

$$E[M_{t+u} \mid I_t] = M_t \tag{24.52}$$

如果我们考虑随机选择时间，那么这种鞅性质还存在吗？

在某些条件下，答案是肯定的．设 τ_1 和 τ_2 为两个独立的停止时间，关于 I_t 可测，且满足：

$$P(\tau_1 < \tau_2) = 1 \tag{24.53}$$

那么，鞅性质仍然成立：

$$E[M_{\tau_2} \mid I_{\tau_1}] = M_{\tau_1} \tag{24.54}$$

在利用等价鞅测度表示资产价格时，该性质显然非常重要．衍生品的执行日期是随机的，但这并不能阻碍等价鞅测度的使用．当 τ 为随机时，随机停止情况下的资产价格仍然是在概率 Q 下的鞅．

24.7.2 Dynkin 公式

设过程 B_t 满足：

$$\mathrm{d}Bt = a(B_t)\mathrm{d}t + \sigma(B_t)\mathrm{d}W_t \tag{24.55}$$

设 $f(B_t)$ 为关于该过程的二阶可微有界函数．

考虑停时 τ 满足：

$$E^Q[\tau] < \infty \tag{24.56}$$

那么我们有

$$E^Q[f(B_\tau) \mid B_0] = f(B_0) + E^Q\left[\int_0^\tau Af(B_s)\mathrm{d}s \mid B_0\right] \tag{24.57}$$

该表达式叫做 Dynkin 公式．它给出了关于停时的函数期望值的简便表达．算子 A 和往常一样，是生成元．

24.8 小结

本章介绍了停时间的概念．该概念在美式衍生品定价以及动态规划中非常实用．我们还介绍了二叉树模型和一类马尔可夫链的密切联系．

24.9 参考阅读

本章有 3 个重要的主题．首先是停时的内容．提早执行是一个最优停止问题．我们推荐 Dynkin 等人(1999)．其中的方法既经典又直观．如果读者想要学习更多关于经典停时问题的内容，可以参考 Shiryayev(1978)．我们提到的第二个主题是动态规划，虽然这对我们来说只是一个边缘问题．有许多关于动态规划的书籍可以参考．

最后一部分内容是停时的数值计算．对于这部分内容，读者可以参考 Broadie 和 Glasserman(1998)．

24.10 习题

1. 一名玩家遇到了以下情况．每个 t 时刻掷一枚硬币($t=1,2,3,\cdots,T$)，设他获得的总奖金为 W_t．他可以决定停止或者继续游戏．若继续，则他在 $t+1$ 时刻抛掷一枚新的硬币，以此类推．问题是，什么时候是最佳停止时间？我们考虑几种不同情况．首先是翻倍或没有的情况．$t=T$ 时刻的总奖金为：

412

$$W_T = \prod_{t=1}^{T}(z_t+1)$$

其中 z_t 是二项随机变量：$z_t=\pm 1$，概率各为$\frac{1}{2}$．

因此在每一个阶段，奖金要么翻倍，要么等于 0.

(a) 根据已知条件，能否计算 T 时刻的期望奖金 $E[W_T]$？

(b) 这个游戏的最佳停止时间是什么时候？

(c) 假设我们增加奖励，对每个阶段的奖金乘以一个大于等于 1 的数．假设现在奖金为：

$$\overline{W_t} = \frac{2n}{n+1}\prod_{t=1}^{T} z_t + 1$$

$T=1,2,3\cdots$证明：若如果我们在 T_k 停止，期望奖金为$\frac{2k}{k+1}$.

(在 T_k 结束表示经历了 k 次投掷．)

(d) 奖金能够达到的最大值是多少？

(e) 是否有最优停止规则？

2. 再次考虑上述问题．假设我们掷一枚硬币 T 次，得到的全部是$+1$．则奖金为

$$W_T = \frac{T(2^T+1)}{T+1}$$

(a) 证明再掷一次的条件期望奖金为：

$$E[W_{T+1} \mid W_T = \omega_T^*] = 2^{T+1}\frac{T+1}{T+2}$$

(b) 该期望与 W_T 相比如何？玩家是否应该“停止”？

(c) 如果玩家在连胜期间一直选择继续游戏，那么他可以持续多久？

(d) 在某一时刻，$z_t=-1$ 的概率是多少？

(e) 你怎么解释这个悖论？

3. 假设你已知以下数据：无风险利率为 6%，股票价格服从 $\mathrm{d}S_t=\mu S_t\mathrm{d}t+\sigma S_t\mathrm{d}W_t$，波动率为 12%，无股利，股票现价为 100. 根据这些数据，你需要估计关于该股票的美式看涨期权现价．期权的执行价格为 100，期限为 200 天．

(a) 选择合适的时间间隔 Δ，使二叉树为 4 步．隐含的 u 和 d 是多少？

(b) 隐含的上升概率是多少？

(c) 确定股票价格 S_t 的二叉树．

(d) 确定看涨期权价格 C_t 的二叉树．

(e) 下面是重要问题：该期权有可能被提前执行吗？

4. 假设上题中讨论的股票支付股利．其余所有参数保持不变．考虑以下 3 种股利支付方式，确定期权是否会被提前执行．

(a) 股票在每个节点支付股利率为 4%的股利．

413

(b) 股票只在第三个节点支付股票价格 5%的股利．其余时间不支付股利．

(c) 股票在第三个节点支付 5 美元的股利．

5. 一名决策者希望用金融工具 k_t 来控制某目标变量 Y_t 的轨迹．该决策者有以下目标函数

$$U=\sum_{i=1}^{4}[2(k_t-k_{t-1})^2+100(Y_t)^2]$$

外部环境对该决策者有以下限制：

$$Y_t=0.2k_t+0.6Y_{t-1}$$

已知初始值 Y_0 为 60.

(a) 在阶段 $t=4$，k_t 的最佳选择是多少？

(b) 在阶段 $t=3$，k_t 的最佳选择是多少？

(c) 根据以上计算，你能否指出规律并说出当 $t=1$ 时 k_t 的最佳选择是多少？

(d) 确定表示 $t=1,2,3,4$ 时最优收益的价值函数 V_t.

(e) 画出价值函数 V_t 并解释其含义．

6. X_1，X_2，…是独立同分布的随机变量且具有有限期望，若 N 为 X_1，X_2，…的停时且 $E(N)<\infty$，证明

$$E\Big(\sum_{n=1}^{N}X_n\Big)=E(N)E(X) \tag{24.58}$$

7. 模型中的股票价格服从以下 SDE：

$$\mathrm{d}S_t=\mu S_t\mathrm{d}t+\sigma S_t\mathrm{d}W_t \tag{24.59}$$

假设模型中无风险利率为 3%，无股利支付．此外，股票现价为 70 美元．在此框架下，写出 1 年期美式看跌期权的定价程序，期权执行价格为 55 美元．在决策树中采用最优执行决策．

414

第 25 章　调整及估值技巧综述

衍生品定价的基础是标的资产(包括股票、商品、利率及汇率)的演化服从随机过程．我们为所研究的模型找出一定的定价程序，从而计算衍生合约的价格．衍生合约定价程序的输入变量包括两类互不相关的参数，分别是(a)合约/市场参数，和(b)模型参数．合约/市场参数反映了衍生品在合约中所规定的特性，包括合约的期限、执行价格、障碍水平等(显然与模型无关).模型参数与演化模型的选择有关，具有主观性，且完全依赖于模型．

所有的定价模型都需要一个参数集来充分定义各种模型的动态变化规律．为了使模型与真实市场相关性更强，更适合用于定价、风险管理或交易，我们必须进行一些校准．校准是确定参数集的过程，它使得流动交易工具的模型价格与市场价格接近吻合．这些流动工具被称为基准或校准工具，它们校准后的价格记录在市场速览中．对于以这些校准工具为基础的模型来说，所谓的校准过程为它们给出了最佳参数集．在大多数应用中，该集合代表了已知市场中大多数(如果不是全部)可流动交易衍生品集合．为保证得到的模型价格接近它们在真实世界中的价格，校准过程通常进行得很频繁(例如每天两到三次)．校准过程是衍生品定价过程中的关键步骤，它也可以用于为流动性更差的衍生品定价，此时需要使用流动性更强的工具的价格．此外，它还可以用于捕捉流动衍生品的套利机会．

当模型参数集大于校准工具价格集时，显然解不唯一．这种问题被称为过参数化．这种情况通常发生在市场上衍生品种类很少但动态变化却很复杂的情况下．过参数化时，无法得到使模型价格与调整工具市场价格精确匹配的参数集．因此在实践中，可以通过解某个特殊构造的最优化问题，得到近似解．反过来，也存在模型类别过窄的情况，此时无法得到足够大的可以与校准工具相匹配的价格集合，则解不存在．这种问题被称为参数化不充分．这种情况最典型的例子就是在 Black-Scholes 框架下的期权定价，该过程中只存在一个自由参数——波动率，而不是很多种流动期权．在参数化不足的情况下，通常使校准工具市场价格的真子集与模型价格在不同模型参数下相匹配．最典型的例子是经典 Black-Scholes 波动率曲面，它使得任意一种流动期权都具有校准后的波动率．一般来说，这就是模型校准不适定的原因(Cont，2010).

25.1　校准公式

我们假设一个基础过程的演化服从某个基础过程，并已知衍生合约的定价步骤之后，就可以设定目标函数来进行校准．目标函数一般为均方误差．我们就可以利用最优化方法使该目标函数最小化．在一般情况下，我们可以将最优化问题写成下面的形式：

$$\inf_{\Theta \in O} \sum_{i=1}^{I} H(C_i^{\Theta} - C_i) \tag{25.1}$$

其中 O 为模型参数空间，H 为目标函数，其自变量为市场价格与模型价格的差 $C_i^{\Theta}-C_i$. 只要我们将该校准问题公式化，最优化算法就可以用于计算解和确定调整后的模型参数．目标函数通常为：

$$H(C_i^{\Theta}-C_i)=w_i\ |C_i^{\Theta}-C_i|^p$$

或

$$H(C_i^{\Theta}-C_i)=w_i\ \left|\frac{C_i^{\Theta}-C_i}{C_i}\right|^p$$

或

$$H(C_i^{\Theta}-C_i)=w_i\ |\ln C_i^{\Theta}-\ln C_i|^p$$

其中 $p\geqslant 1$，w_i 为正权重，通常与 C_i 买卖差价的平方成反比．目标函数不同，结果也不同. 校准问题最常见的一种公式是最小平方公式：

$$\inf_{\Theta\in O}\sum_{i=1}^{1}w_i(C_i^{\Theta}-C_i)^2 \tag{25.2}$$

416

25.2 基础模型

我们将涉及 3 个模型，即 Black-Scholes 模型、局部波动率模型及方差伽马模型．

25.2.1 几何布朗运动——Black-Scholes 模型

第 12 章中提到过 Black-Scholes 模型．对于依赖于终端条件和边界条件的衍生证券，Black-Scholes PDE 给出了它们的价格．考虑 Black-Scholes PDE

$$\frac{\partial V}{\partial t}+\frac{1}{2}\sigma^2S^2\ \frac{\partial^2 V}{\partial S^2}+(r-q)S\ \frac{\partial V}{\partial S}=rV \tag{25.3}$$

假设终端条件为期限为 T 的看跌期权收益

$$V(S,T)=\max(K-S,0) \tag{25.4}$$

该 PDE 在 $t<T$ 时的解析解为欧式看跌期权的 Black-Scholes 期权定价公式，即

$$V(S,t)=Ke^{-r(T-t)}\Phi(-d_2)-Se^{-q(T-t)}\Phi(-d_1) \tag{25.5}$$

其中

$$d_1=\frac{\ln(S/K)+(r-q+\sigma^2/2)(T-t)}{\sigma\sqrt{T-t}} \tag{25.6}$$

$$d_2=d_1-\sigma\sqrt{T-t} \tag{25.7}$$

该模型中唯一无法观测的模型参数是基础资产的波动率．它是影响期权价值的唯一变量．为找出期权价格的最小可能值和最大可能值，我们对波动率变大(假设趋于无穷大)和波动率趋于 0 的情况进行计算．

情况 1：波动率趋于无穷大($\sigma\to\infty$)

根据(25.6)及(25.7)，$d_1\to\infty$，$d_2\to\infty$. 因此 $\Phi(-d_1)=\Phi(-\infty)=0$，$\Phi(-d_1)=\Phi(\infty)=1$，因此 $V(S,t)=Ke^{-r(T-t)}$.

情况 2：波动率趋于 0($\sigma \to 0$)

这种情况取决于 S 落在执行价格 K 的左边还是右边．

(a) 如果 $S>K$，根据(25.6)及(25.7)可以看出 d_1 和 d_2 都趋于无穷，即 $\Phi(-d_1)=\Phi(-d_2)=\Phi(-\infty)=0$，因此 $V(S,t)=0$.

(b) 如果 $S<K$，根据(25.6)及(25.7)可以看出 d_1 和 d_2 都趋于负无穷，即 $\Phi(-d_1)=\Phi(-d_2)=\Phi(\infty)=1$，因此 $V(S,t)=K\mathrm{e}^{-r(T-t)}-S\mathrm{e}^{-q(T-t)}$.

无论基础过程是怎样的，上述情况都不受影响，因此它是不受模型影响的．我们可以用二分法、Newton-Raphson 法等方法找出波动率，使其模型价格与已知看跌期权费相匹配，只要看跌期权费在 0 到 $V(S,t)=K\mathrm{e}^{-r(T-t)}$之间．通过这种方式得到的波动率称为隐含波动率．

在现实中，市场上存在多种期权费，用一种波动率就可以代表整个波动率面的假设是不现实的，因此不能使用 Black-Scholes 方法．虽然如此，市场从业者仍在 Black-Scholes 隐含波动率的基础上对期权价进行报价．

在调整过程中我们需要确定的参数集是 $\Theta=\{\sigma\}$，可交易期权的隐含波动率．通过表 25-1和 25-2，我们可以得到每种期权的隐含波动率，分别形成看涨期权费和看跌期权费．图 25-1a 及 25.1b 分别画出了看涨期权和看跌期权的隐含波动率面．

表 25-1 S&P 500 看涨期权费(2012.3.27)

期限/天	执行价格	Bid	Ask	r	q	远期合约价格
24	1405	21.7	23.0	0.2324	1.5314	1410.55
	1410	19.1	20.0	0.2324	1.5314	1410.80
	1415	15.9	17.1	0.2324	1.5314	1410.55
	1420	13.6	14.6	0.2324	1.5314	1410.70
	1425	11.0	12.2	0.2324	1.5314	1410.55
52	1405	30.8	32.6	0.3240	2.1087	1407.85
	1410	28.2	29.6	0.3240	2.1087	1408.00
	1415	25.5	26.8	0.3240	2.1087	1408.00
	1420	22.8	24.1	0.3240	2.1087	1407.95
	1425	20.6	21.6	0.3240	2.1087	1408.09
80	1405	39.5	41.4	0.4331	2.0990	1406.30
	1410	36.8	38.3	0.4431	2.0990	1406.50
	1415	33.8	35.6	0.4331	2.0990	1406.29
	1420	31.3	32.7	0.4331	0.0990	1406.54
	1425	28.4	30.1	0.4331	2.0990	1406.38
93	1375	61.8	66.3	0.4826	2.0819	1405.74
	1400	47.0	48.4	0.4826	2.0819	1405.81
	1425	32.8	34.1	0.4826	2.0819	1405.83
	1450	21.2	22.5	0.4826	2.0819	1405.04
	1475	12.6	13.9	0.4826	2.0819	1405.66
115	1375	67.4	69.7	0.5572	1.9816	1404.05

（续）

期限/天	执行价格	Bid	Ask	r	q	远期合约价格
	1400	51.4	53.6	0.5572	1.9816	1404.41
	1425	37.3	39.2	0.5572	1.9816	1404.21
	1450	25.5	27.2	0.5572	1.9816	1404.27
	1475	16.2	17.7	0.5572	1.9816	1404.12
178	1375	80.7	83.2	0.7316	2.0899	1400.04
	1400	65.2	67.3	0.7316	2.0899	1400.45
	1425	50.7	53.1	0.7316	2.0899	1400.31
	1450	39.0	40.5	0.7316	2.0899	1400.22
	1475	28.0	29.7	0.7316	2.0899	1400.33
184	1375	82.2	84.5	0.7436	2.0980	1399.14
	1400	66.6	68.8	0.7436	2.0980	1399.15
	1425	52.5	54.5	0.7436	2.0980	1399.00
	1450	40.1	41.9	0.7436	2.0980	1399.86
	1475	29.5	31.0	0.7436	2.0980	1399.71
269	1375	95.5	98.5	0.8831	2.1546	1393.82
	1400	80.4	83.4	0.8831	2.1546	1393.31
	1425	66.5	69.1	0.8831	2.1546	1393.59
	1450	53.7	56.4	0.8831	2.1546	1393.63
	1475	42.4	44.8	0.8831	2.1546	1393.52

417~418

表 25-2　S&P 500 看跌期权费(2012.3.27)

期限/天	执行价格	Bid	Ask	r	q	远期合约价格
24	1405	16.2	17.4	0.2324	1.5314	1410.55
	1410	18.1	19.4	0.2324	1.5314	1410.80
	1415	20.3	21.6	0.2324	1.5314	1410.55
	1420	22.7	24.1	0.2324	1.5314	1410.70
	1425	25.3	26.8	0.2324	1.5314	1410.55
52	1405	27.9	29.8	0.3240	2.1087	1407.85
	1410	29.9	31.9	0.3240	2.1087	1408.00
	1415	32.1	34.2	0.3240	2.1087	1408.00
	1420	34.4	36.6	0.3240	2.1087	1407.94
	1425	36.9	39.1	0.3240	2.1087	1408.09
80	1405	38.1	40.2	0.4331	2.0990	1406.30
	1410	40.2	41.9	0.4431	2.0990	1406.50
	1415	42.3	44.5	0.4331	2.0990	1406.29
	1420	44.6	46.3	0.4331	0.0990	1406.54
	1425	47.0	48.7	0.4331	2.0990	1406.38
93	1375	32.6	34.1	0.4826	2.0819	1405.74
	1400	41.0	42.8	0.4826	2.0819	1405.81
	1425	51.6	53.6	0.4826	2.0819	1405.83
	1450	64.8	68.7	0.4826	2.0819	1405.04
	1475	80.1	84.9	0.4826	2.0819	1405.66
115	1375	38.5	40.6	0.5572	1.9816	1404.05

(续)

期限/天	执行价格	Bid	Ask	r	q	远期合约价格
	1400	47.2	49.0	0.5572	1.9816	1404.41
	1425	57.8	60.2	0.5572	1.9816	1404.21
	1450	70.6	73.4	0.5572	1.9816	1404.27
	1475	86.2	89.2	0.5572	1.9816	1404.12
178	1375	55.7	58.3	0.7316	2.0899	1400.04
	1400	64.8	66.8	0.7316	2.0899	1400.45
	1425	75.4	77.6	0.7316	2.0899	1400.31
	1450	87.7	91.0	0.7316	2.0899	1400.22
	1475	102.0	104.5	0.7316	2.0899	1400.33
184	1375	57.9	60.7	0.7436	2.0980	1399.14
	1400	67.0	70.1	0.7436	2.0980	1399.15
	1425	77.7	81.1	0.7436	2.0980	1399.00
	1450	88.5	93.4	0.7436	2.0980	1399.86
	1475	102.3	108.2	0.7436	2.0980	1399.71
269	1375	77.1	79.5	0.8831	2.1546	1393.82
	1400	86.8	90.3	0.8831	2.1546	1393.31
	1425	97.6	100.4	0.8831	2.1546	1393.59
	1450	109.5	112.6	0.8831	2.1546	1393.63
	1475	122.9	126.2	0.8831	2.1546	1393.52

419

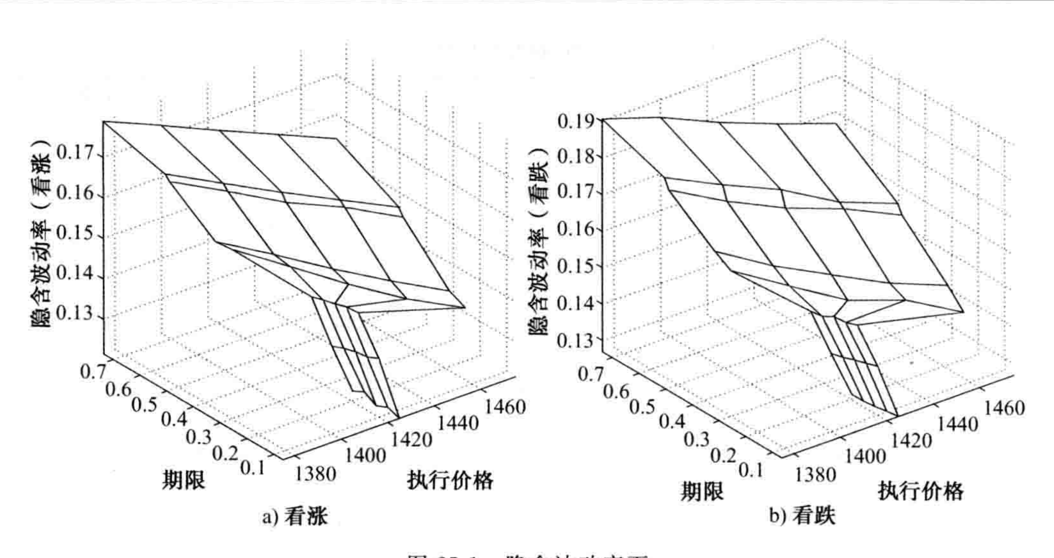


图 25-1　隐含波动率面

25.2.2　局部波动率模型

Black-Scholes 模型被公认为是最简单的衍生品定价公式，可用于对许多简单的衍生合约定价．然而，由于该模型需要波动率曲面(这意味着不同期权的参数可能不同)以及无法正确地复制标的资产的变化，它在许多奇异衍生品合约的定价和对冲上受到了限制．因而

人们对该模型进行了进一步研究．

为了克服 Black-Scholes 模型的不充分性和局限性，Derman 和 Kani(1994)提出了局部波动率模型，利用当前标的价格和日历时间对波动率参数化．该设计能够使衍生品的定价和对冲与波动率微笑相容，而这已经成为近十多年来一个重要的研究领域．而局部波动率的发展是该领域研究中关键的一步．

当对标准 Black-Scholes 模型参数的拓展时，我们将标的价格和时间都考虑进来，这使得我们可以同时对许多甚至所有流动普通期权合约校准．此外，该模型在对特种衍生品使用时可行性更高．局部波动率模型满足以下随机微分方程：

$$dS_t = (r(t) - q(t))S_t dt + \sigma(S_t, t)S_t dW_t \tag{25.8}$$

这里假设利率和红利都有确定性的期限结构．该模型中的期权价格满足 Black-Scholes 型方程，也就是我们说的一般 Black-Scholes PDE，它能够根据终止条件和边界条件，给出衍生证券的价格．

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2(S,t)S^2\frac{\partial^2 V}{\partial S^2} + (r(t) - q(t))S\frac{\partial V}{\partial S} = r(t)V(S,t) \tag{25.9}$$

局部波动率模型的一大优点是简洁．要得到局部波动率模型的解和衍生品价格，我们只需要对 Black-Scholes 模型进行几个简单的调整．此外，它具有很好的灵活性，能对所有期权定价．因此大多数交易者和公司都经常使用该模型．在该模型下，衍生品的评价和定价非常简单，而对冲保值比率的计算也很直接．此外，根据 Hirsa 等人(2001)的介绍，局部波动率模型重新处理了普通期权的纯跳跃模型．但是，对于路径依赖型期权，局部波动率模型和纯跳跃模型的表现可能差异非常大(Hirsa 等人，2003)．

420

在局部波动率模型中，校准参数的集合为 $\Theta=\{\sigma(S_t,t)\}$．合约参数和确定性市场参数的集合为 $\Lambda=\{S_0,K,T,r(t),q(t)\}$．

局部波动率模型中校准过程的目标是找到局部波动率曲面 $\sigma(S,t)$，使得不同执行价格和期限的期权模型价格与相应的市场价格吻合．完成该目标的一个方法是对校准问题进行如下公式化：

$$\arg\min_{\Theta} \frac{1}{M}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{M} \| V_{ij} - \hat{V}_{ij} \| \tag{25.10}$$

其中 V_{ij} 和 $\hat{V}_{ij}$ 分别是市场和模型价格，对应的执行价格为 K_i，期限为 T_j．对于每个执行价格和期限的组合，我们都假设同一个波动率曲面 $\sigma(S,t)$．而对于每个期限和执行价格的组合，我们要分别求 Black-Scholes PDE，从而求出各期权的模型价格．对于固定的执行价格，支付额相同，但期限是不同的，因而期权价格不同．而固定期限的期权，其支付额根据执行价格有所不同，期权价格也是不同的．

25.2.3 欧式期权的向前偏微分方程

和上一节提到的一样，在使用一般 Black-Scholes PDE 时，我们需要构建局部波动率曲面的校准问题，我们需要求解 PDE，对于每个要校准的衍生品价格和优化过程都是如此．这将

涉及许多 PDE 的求解．如果可能的话，我们希望同时求解不同执行价格和期限的期权价格，曲面上的格点对应了这些执行价格和期限．这样做能够大大减少校准过程的计算时间．

二十世纪九十年代中期发生了一项突破，人们发现在某些模型中，欧式期权价值服从向前演化方程，其中独立变量是期权执行价格和期限．具体而言，Dupire(1994)证明了在持有成本确定以及标的价格服从扩散过程的假设下，无套利意味着欧式期权价格满足某个偏微分方程，即现在我们所称的 Dupire 方程．如果假设我们可以观察到任意执行价格和期限的欧式期权价格，那么该向前 PDE 可以用于具体确定标的资产的瞬时波动率函数，其中函数的自变量为标的资产的价格以及时间．

一旦该波动率函数已知，欧式、美式及许多特种期权也可以相继确定，许多常规方法都可以做到这一点．由于这些价值函数能够将衍生工具价格与标的资产价格和时间联系起来，它也可以帮助我们确定许多(由希腊字母表示的)对冲参数．除了可以用于确定波动率函数，向前方程还有其他作用．只要波动率函数已知，无论是具体形式还是先验校准的形式，我们都可以通过有限差分法求解 PDE，从而有效地对一系列相同标的资产、不同执行价格和期限的期权估值．假设已知一个波动率曲面以及格点的位置，这使我们能够同时求解所有格点上的期权价格(对应不同的执行价格和期限)，用同样的方法还能够求出所有满足这些方程的希腊字母. 421 Andreasen(1998)介绍了这些内容．这使我们能够以更低的计算量解决隐含波动率校准问题．

前面提到，Dupire 方程能够给出不同执行价格和期限的欧式看涨期权价格，该方程如下：

$$-\frac{\partial C}{\partial T}+\frac{1}{2}\sigma^2(K,T)K^2\frac{\partial^2 C}{\partial K^2}-(r(T)-q(T))K\frac{\partial C}{\partial K}=q(T)C \tag{25.11}$$

只要知道看涨期权曲面，就能够直接求出局部波动率曲面．观察该 PDE 容易发现，如果假设期权市场价格 $C(K_i, T_j)$已知，我们能够利用以下反推公式具体求出波动率曲面：

$$\sigma(K,T)=\left(\frac{\frac{\partial C}{\partial T}+(r(T)-q(T))K\frac{\partial C}{\partial K}+q(T)C}{\frac{1}{2}K^2\frac{\partial^2 C}{\partial K^2}}\right)^{1/2} \tag{25.12}$$

为了可以使用来自 Dupire 的反推公式，我们必须事先假设看涨期权价格关于执行价格和期限的曲面 $C(K,T)$足够光滑．此外，我们还必须能够计算日历价差$\frac{\partial C}{\partial T}$，蝶形价差$\frac{\partial C}{\partial K}$以及关于执行价格的二阶偏导数$\frac{\partial^2 C}{\partial K^2}$，对看涨期权价格曲面上的每个格点都要计算．

在实践中，这是使用该方法最大的难点．期权市场价格只在有限的格点上已知，而中间的执行价格与期限所对应的期权价格必须通过插值得到．这里使用插值后，有可能会在执行价格和期限的范围内出现静态套利，这不满足不存在静态套利的假设．即使达到了该要求，插值可能使得许多地方变得不可微，从而使局部波动率曲面变得形状古怪，可信度低．对期权价格曲面进行适当地插值，使之与观察到的市场价格相一致，是用相容的期权定价模型对市场价格进行公式化和估值的必要任务．根据 Cont(2010)介绍的内容，局部波动率可能对输入变量的微小变动非常敏感．在下一节对局部波动率曲面校准的介绍中，我

们将探索一些不同的方法来构建局部波动率曲面．是否能够生成光滑、稳定、相容的局部波动率曲面还有待观察．

但是，没有足够的期权价格可供我们构建光滑的看涨期权价格曲面，因而也就不能使用反推公式，那么怎么办呢？首先想到的方法是用外推插值建立看涨期权价格的光滑曲面（例如双三次样条插值），或者假设某些局部波动率的函数结构（5 或 6 个变量），通过优化问题找到合适的参数．注意在第一个方法中无需用到优化或者目标函数．

利用表 25-1 中的看涨期权价格，我们能够使用双三次样条插值，得到比当前网络更细化的价格网络．然后利用(25.12)，可得局部波动率曲面．所得的局部波动率曲面如图 25-2 所示．从图中可以看到，该曲面非常不光滑．由于局部波动率曲面像正弦双曲函数，我们假设该曲面服从以下函数．

$$\sigma(K,T)=\frac{1}{2}(c_1\exp(K/c_2)+c_3\exp(-K/c_2))\exp(-\beta T) \tag{25.13}$$

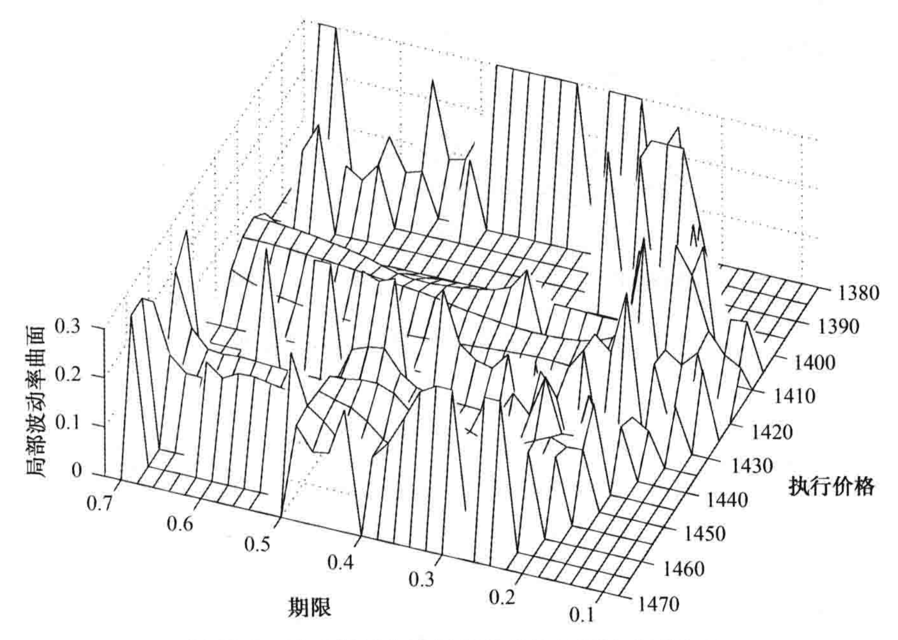


图 25-2　看涨期权价格插值得到的局部波动率曲面

通过校准我们为局部波动率函数找到了以下参数

422

$$c_1=-0.33 \tag{25.14}$$

$$c_2=1404.27 \tag{25.15}$$

$$c_3=2.19 \tag{25.16}$$

$$c_4=2222.02 \tag{25.17}$$

$$\beta=-1.13 \tag{25.18}$$

我们将参数代入函数(25.13)，即可得到图 25-3．我们没有对 β 的符号施加限制．可以看

到，波动率随着期限的增加而增加，这与我们想象的有所不同．图 25-4 中我们画出了看涨期权价格曲面，并与市场价格对比．可以看到两者紧密贴合，价格的平均差异为 64.5 美分．这相比于买卖差价非常紧密．但这并不意味由此得到的局部波动率是可接受的．对于其他模型，如 Heston 随机波动率模型或者方差伽马模型，我们需要对衍生工具的定价模型进行校准．我们建立目标函数，利用优化程序找到使目标函数最小的最优参数．

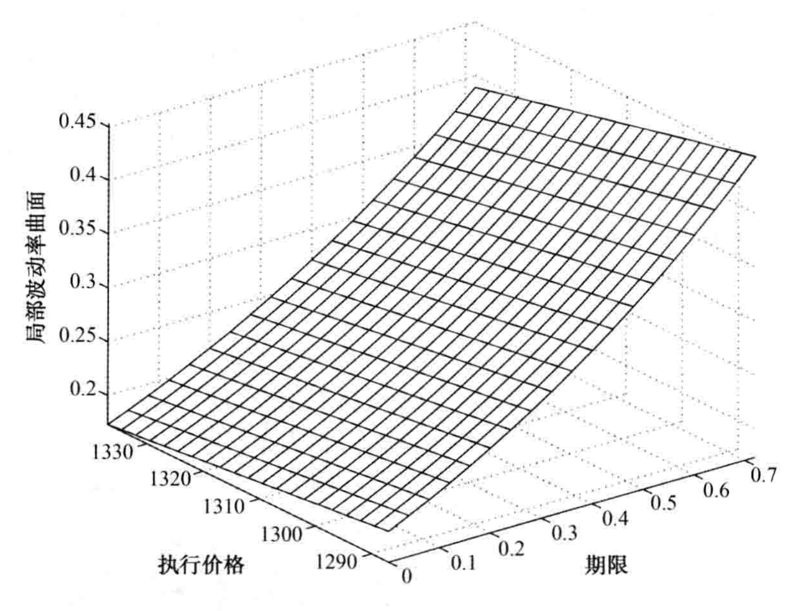


图 25-3　由一函数确定的局部波动率曲面

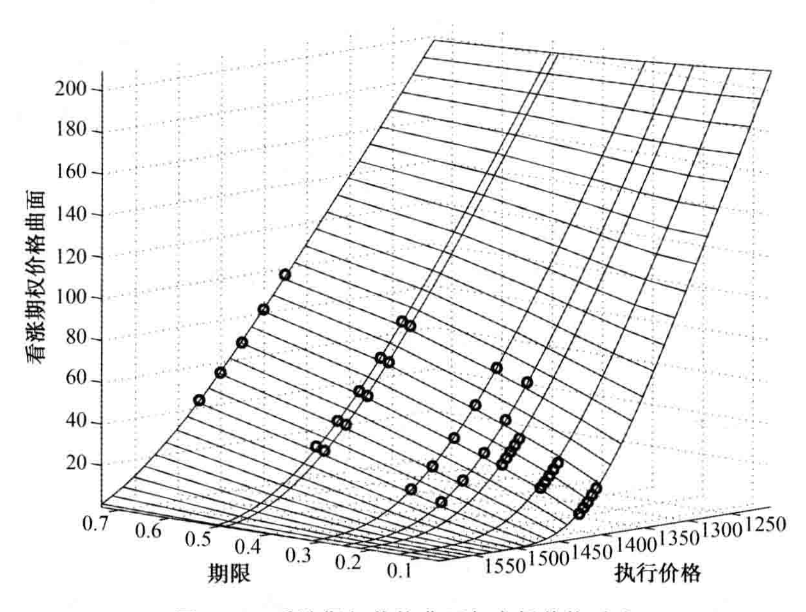


图 25-4　看涨期权价格曲面与市场价格对比

在 Hirsa 等人(2003)中，作者提出，根据所考虑的模型计算执行价格和期限在一定范围内的看涨期权价格，继而对参数校准，并将价格带入方程(25.12)，然后计算出相应的波动率曲面．任何随机波动率模型，如 Heston 随机波动率模型，都能够关于所有执行价格和期限，对市场上的欧式期权价格进行校准．这些校准后的模型能够生成一个光滑的看涨/看跌价格曲面．对于纯跳跃框架下的随机波动率模型，该方法能够生成非常光滑的局部波动率曲面．所得的局部波动率曲面完全依赖于模型中欧式期权价格的插值情况．显然该曲面不是唯一的．他们分别使用 VGSA 和 Heston 随机波动率模型，用该方法给出了 S&P500 股票指数的局部波动率曲面．

423～424

25.2.4 方差伽马模型

期权价格中有时存在幅度较大的波动率微笑，尤其是在期限较短的期权中．这会使得扩散模型中的模型波动率产生很大变化．这造成了非常数波动率的存在，如局部波动率、常数弹性波动率、Heston 随机波动率模型等．波动率微笑的存在使许多研究者都使用跳跃模型对标的资产建模．这类模型允许资产价格发生离散跳跃，能较好地解释波动率微笑．方差伽马模型就是此类模型之一，我们曾在 11.6.2 节有过介绍．我们介绍过，它是一个三参数模型，校准参数为 $\Theta=\{\sigma,\theta,\nu\}$，分别是波动率、偏度和峰度．

结果显示，对于一个固定的期限，该模型确实能够较好地适应不同的执行价格．这意味着对每个期限我们要单独进行校准，并得到单独的参数集合．在校准时，我们主要观察方差伽马模型中参数关于期限的变化情况．其中波动率参数关于期限是递减的：期限越长，波动率参数越小，但是总的来说波动率集中在一个较小的范围内．峰度关于期限是递增的；而偏度随着期限递减．在股份期权中我们通常会观察到负偏态(向右偏)的情况．

这里我们展示两个方差伽马模型中校准的例子．在第一个例子中，我们可以观察方差伽马模型的极限行为．我们将看到，如果虚值看涨或看跌期权以常波动率的 Black-Scholes 模型价格出现在市场上，方差伽马模型能够侦测出该价格来自于常波动率的 Black-Scholes 模型．这意味着在 $\nu=0$ 和 $\theta=0$ 的特殊情况下，方差伽马模型能够将纯扩散模型复原．

本例中的参数为现价 100 美元，波动率 $\sigma=0.40$，期限 $T=0.5$ 年，无风险利率 $r=2\%$，红利率 $q=1.0\%$，执行价格为 $80,85,90,\cdots,125$. 方差伽马模型的初始参数为 $\sigma=0.10$，$\nu=0.20$ 以及 $\theta=-0.20$. 用 Nelder-Mead 优化方法校准之后的参数为 $\sigma=0.403\,928$，$\nu=0.021\,429$，$\theta=-0.049\,330$. 在图 25-5 中我们画出了方差伽马校准价格和 Black-Scholes 价格．显然，它们贴合非常紧密，很难区分出来．

在第二个校准的例子中，我们利用 S&P 500 期权校准，得到了方差伽马参数．表 25-3 展示了利用 2012 年 3 月 27 日的不同期限的 S&P 500 虚值欧式期权(看涨/看跌)校准得到的方差伽马参数．对这些方差伽马参数的初始预测是 $\sigma=0.20$，$v=0.20$，$\theta=-0.20$. S&P 500 在 2012 年 3 月 27 日的即期价格为 1412.52 美元．

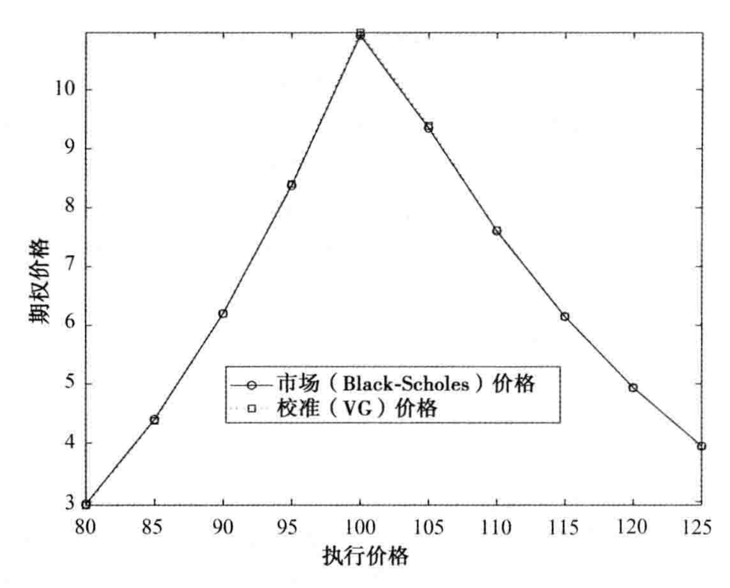


图 25-5 VG 对 Black-Scholes 的校准

表 25-3 S&P 500 校准的 VG 参数(2012.3.27)

期限	σ	ν	θ	r	q	即时
0.065 75	0.1230	0.0425	−0.3398	0.002 32	0.015 31	1412.52
0.142 47	0.1494	0.1567	−0.1512	0.003 24	0.021 09	1412.52
0.219 18	0.1432	0.2306	−0.1952	0.004 33	0.020 99	1412.52
0.254 79	0.1438	0.2333	−0.2082	0.004 83	0.020 82	1412.52
0.315 07	0.1513	0.3385	−0.1727	0.005 57	0.019 82	1412.52
0.487 67	0.1649	0.6021	−0.1460	0.007 32	0.020 90	1412.52
0.504 11	0.1659	0.6344	−0.1487	0.007 44	0.020 98	1412.52
0.736 99	0.1678	0.8594	−0.1430	0.008 83	0.021 55	1412.52

我们利用表 25-1 和表 25-2 的期权数据分别得到关于不同期限 S&P 500 期权校准的方差伽马模型参数．利用校准后的方差伽马模型，我们生成看涨和看跌期权价格，将它们画出并与市场数据作对比．我们会发现取值为负数的 θ 总的来说随着期限下降．这一单调性反映了风险厌恶心理所造成的负偏度，而偏度在隐含年化风险中性密度下递减．参数 ν 随着期限的增加持续增长，反映隐含年化风险中性密度下超值峰度的增加．波动率参数随着期限的增加而增加，但变化幅度相对较小．为便于说明，我们分别用图 25-6 和图 25-7 给出 S&P 500 在 2012 年 3 月 27 日期限为 24 天和 93 天的期权价格，其中包括方差伽马价格和市场价格(中间价)．在这些图中，右侧图表示看涨期权价格，左侧图表示看跌期权价格．我们还给出了出价和询价的误差区间．在校准时我们使用 Nelder-Mead 优化方法．假如我们使用网格搜索法，会出现什么情况呢？假设我们已经画好网格并具有以下 3 种参数.

425

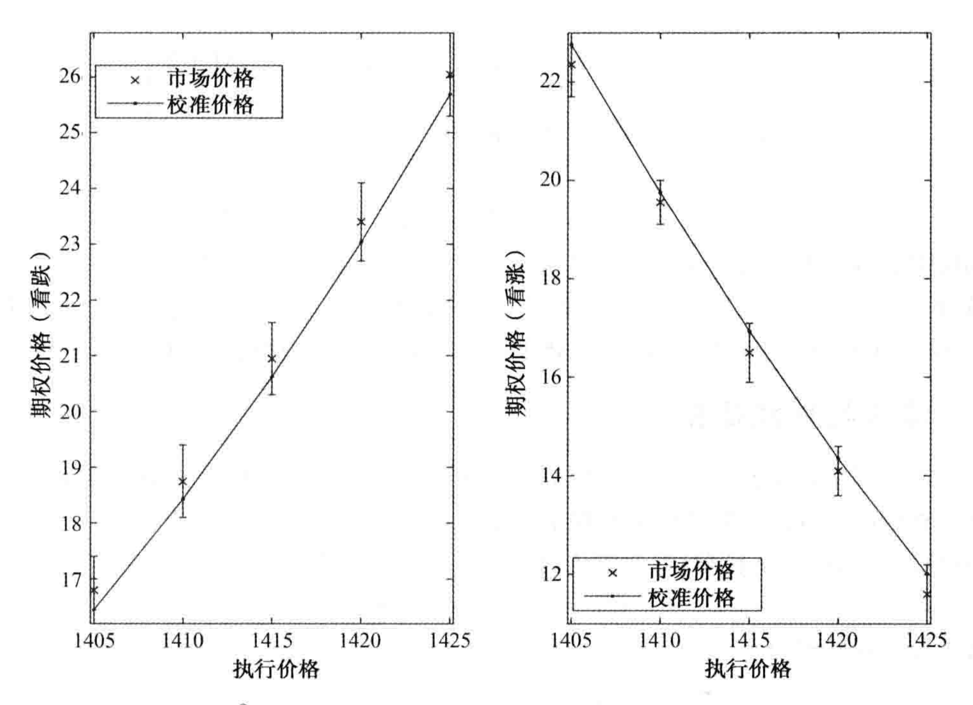


图 25-6　S&P 500 的 24 天期限期权 VG 和市场价格对比(2012.3.27)

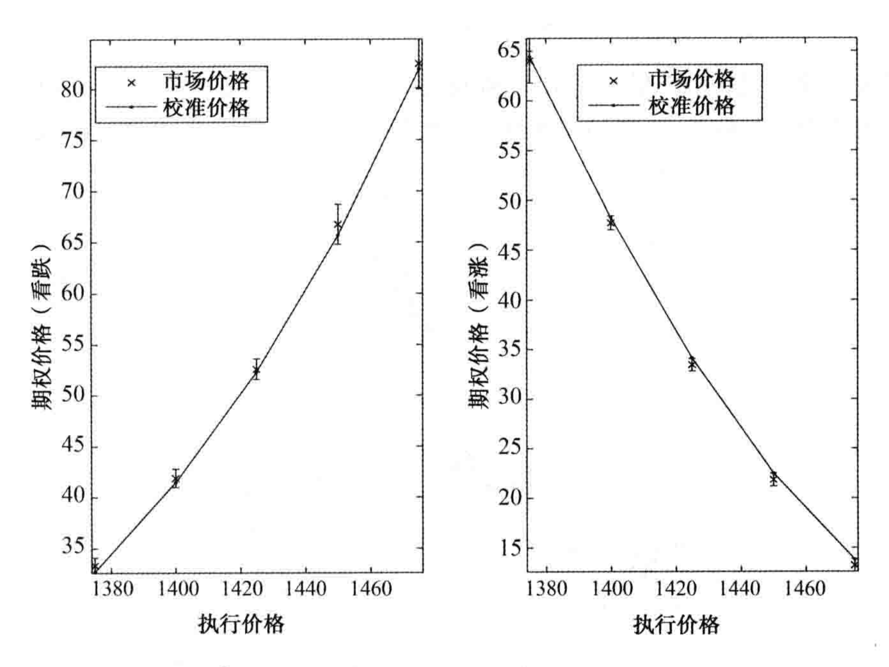


图 25-7　S&P 的 93 天期限期权 VG 和市场价格对比(2012.3.27)

$$\overline{D}=\begin{Bmatrix}\sigma_i=a+(i-1)\Delta\sigma;\Delta\sigma=\dfrac{b-a}{M};i=1,\cdots,M+1\\\nu_j=c+(j-1)\Delta\nu;\Delta\nu=\dfrac{d-c}{N};j=1,\cdots,N+1\\\theta_k=e+(k-1)\Delta\theta;\Delta\theta=\dfrac{f-e}{K};k=1,\cdots,K+1\end{Bmatrix}\tag{25.19}$$

同时考虑以下上下限：对于波动率 $a=0.05$，$b=0.40$；对于峰度 $c=0.01$，$d=0.90$；426 对于偏度 $e=-0.40$，$f=0.20$，格点数为 $M=15$，$N=89$，$K=60$. 对于每个期限我们有 10 个期权(5 个看涨 5 个看跌)，我们需要进行 878 400 次期权估值，计算量非常大.

25.3 滤波与估测概括

在校准过程中，我们并没有用到价格的时间序列，而是只用到了价格的市场速览. 可能会有这种情况，我们希望所得的参数集适用于与历史价格，而非市场速览. 我们以几何布朗运动(Black-Scholes 模型)的参数估测为例. 考虑以下随机微分方程

$$\mathrm{d}S_t=\mu S_t\mathrm{d}t+\sigma S_t\mathrm{d}W_t\tag{25.20}$$

该随机微分方程的解为

$$S_t=S_0\exp\left(\mu-\frac{\sigma^2}{2}\right)t+\sigma W_t\tag{25.21}$$

$$=S_0\exp\left(\mu-\frac{\sigma^2}{2}\right)t+\sigma\sqrt{t}Z\tag{25.22}$$

其中 Z 服从标准正态分布 $\mathcal{N}(0,1)$. 假设漂移率为 5%，$\mu=0.05$，波动率为 20%，$\sigma=0.25$，可模拟出为期 1000 天的样本路径，如图 25-8. 拟合从 S_0 出发，满足

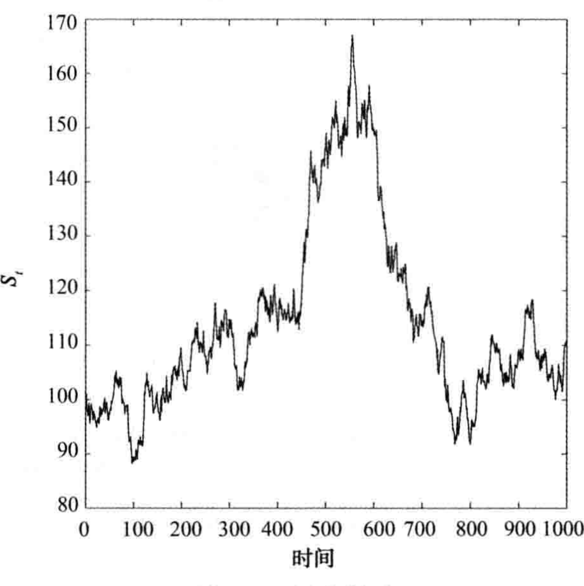


图 25-8　网格例子

$$S_{t+1} = S_t \exp\left(\mu - \frac{\sigma^2}{2}\right)\Delta t + \sigma \sqrt{\Delta t} Z \tag{25.23}$$

极大似然的方法有许多种．这里我们采用(a)网格搜索法(brute-force)，(b)单纯形法(Nelder-Mead)．首先，我们使用网格搜索法(brute-force)进行极大似然．设定网格为 427

$$\overline{D} = \begin{Bmatrix} \mu_j = a + (j-1)\Delta\mu; \Delta\mu = \dfrac{b-a}{N}; j = 1, \cdots, N+1 \\ \sigma_k = c + (k-1)\Delta\sigma; \Delta\sigma = \dfrac{d-c}{M}; k = 1, \cdots, M+1 \end{Bmatrix} \tag{25.24}$$

其中对于漂移率的上下限我们取 $a=0.01$，$b=0.20$；对波动率的上下限我们取 $c=0.05$，$d=0.50$．同时假设漂移率和波动率的空间步数分别为 $N=190$，$M=45$．这意味着步长为 $\Delta\mu=0.001, \Delta\sigma=0.01$．对于几何布朗运动，$t$ 时刻的似然函数为

$$f(S_{t+1}, \mu, \sigma \mid S_t) = \frac{1}{\sqrt{2\pi}} \frac{1}{S_{t+1}\sigma\sqrt{\Delta t}} \times \exp\left(-\frac{(\log S_{t+1} - \log S_t - \left(\mu - \frac{\sigma^2}{2}\right)\Delta t)^2}{2\sigma^2 \Delta t}\right) \tag{25.25}$$

因此，总(对数)似然函数为

$$\sum_{i=1}^{N} \log f(S_i, \mu, \sigma \mid S_{i-1}) = \sum_{i=1}^{N} -\log\left(S_i\sigma\sqrt{\Delta t} - \frac{(\log S_i - \log S_{i-1} - \left(\mu - \frac{\sigma^2}{2}\right)\Delta t)^2}{2\sigma^2 \Delta t}\right) \tag{25.26}$$

通过网格搜索我们可知，取最大值时 $\mu=0.045$ 和 $\sigma=0.21$，与 μ 和 σ 的真实值非常接近．

现在使用 Nelder-Mead 优化方法，设初始值为$[\hat{\mu}, \hat{\sigma}]=[0.01, 0.01]$，最大化对数似然函数(最小化负的对数似然函数)可得 $\mu=0.0446$ 和 $\sigma=0.2065$，与网格搜索法的结果以及真实值都非常接近．一般来说，估计波动率比估计漂移率容易得多．参见本章最后的例子．对于其他项，较宽的步长效果会更好，但可能可行性较低． 428

以上估测的例子基于极大似然函数．在该例中，极大似然函数以解析式的形式存在．下面我们考虑一个不是以解析形式存在的例子．假设某状态服从以下简单线性模型：

$$x_{t+1} = a x_t + \lambda z_{t+1} \tag{25.27}$$

其中 $z_{t+1} \sim \mathcal{N}(0, 1)$．令 y_{t+1} 为 $t+1$ 时刻的市场价格(观察价格)，假设它与模型价格具有以下关系：

$$y_{t+1} = h(x_{t+1}; \Theta) + \sigma u_{t+1} \tag{25.28}$$

其中 $u(t+1) \sim \mathcal{N}(0, 1)$，$h$ 是定价模型，Θ 参数集合，包括 a，λ，σ．已知观察值 y_{t+1}，$t=1,\cdots,T$ 的时间序列，我们不能计算出似然函数的解析形式，这是由于 x_{t+1} 是未知变量．以状态 x_t 为条件，我们可以找到(条件)似然函数；然而，有许多不同的 x_{t+1} 我们都可以作为条件．状态 x_{t+1} 叫做隐藏状态，不能够直接被观察到．

有两种可能的情况．第一种，参数集合 Θ 已知，我们通过观察 y_{t+1} 找到 x_{t+1} 的最佳估计值．从 $t=0$ 开始，我们对 x_0 取任意值．已知 y_1，我们可求出 x_1 的最佳估计．在每个时

429 间段重复该步骤．通过这些步骤，我们可以建立隐藏状态的时间序列．注意到，我们可以利用根据过去时间步得到的状态估计值以及当前时间步的观察值，来计算当前时间状态的最佳估计．因此可以假设 t 时刻的状态 x_t 已知．现在，已知 $t+1$ 时刻的观察值，我们希望计算 x_{t+1} 的最佳估计，即 $\hat{x}_{t+1}$，

$$\hat{x}_{t+1} = E(x_{t+1} \mid y_{t+1})$$

其中 y_{t+1} 是 $t+1$ 时刻的观察值．假设模型价格由 $h(x_{t+1}; \Theta)$ 给出，如前所述，Θ 是参数集．我们假设 $t+1$ 时刻的市场价格（观察值）y_{t+1} 与模型价格有以下关系：

$$y_{t+1} = h(x_{t+1};\Theta) + \sigma u_{t+1} \tag{25.29}$$

其中 $u_{t+1} \sim \mathcal{N}(0,1)$．$\lambda$，$\sigma$ 都是参数集 Θ 中的参数，因而都是已知或已有估计值的．

已知隐藏状态 x_{t+1} 的演化方程，我们首先可以生成 x_{t+1} 的 m 个样本．

$$x_{t+1}^{(i)} = ax_t + \lambda\,\mathcal{N}(0,1) \tag{25.30}$$

其中，$i=1,\cdots,m$．已知 m 个样本 $x_{t+1}^{(i)}$，我们可以计算模型价格的 m 个样本，即 $h(x_{t+1}^{(i)};\Theta)$．然后在已观察到 $t+1$ 时刻市场价格的情况下，我们可以生成 u_{t+1} 的 m 个样本：

$$u_{t+1}^{(i)} = y_{t+1} - h(x_{t+1}^{(i)};\Theta) \tag{25.31}$$

我们定义 $\mathcal{L}^{(i)}$ 为（条件）似然函数

$$\mathcal{L}^{(i)} \equiv \text{似然}(u_{t+1}^{(i)} \mid x_{t+1}^{(i)}) \tag{25.32}$$

这里$\mathcal{L}^{(i)}$ 为

$$\mathcal{L}^{(i)} = \frac{e^{-\frac{\left(u_{t+1}^{(i)}\right)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} \tag{25.33}$$

因此

$$\hat{x}_{t+1} = E(x_{t+1} \mid z_{t+1}) \tag{25.34}$$

$$= \frac{\sum_{i=1}^{M} \mathcal{L}^{(i)} \times x_{t+1}^{(i)}}{\sum_{i=1}^{M} \mathcal{L}^{(i)}} \tag{25.35}$$

这是 x_{t+1} 的最佳估计．要得到下一时刻的最佳估计，需要把我们刚才得到的值作为当前时刻的最佳估计，以此类推．首先可以发现在滤波中我们固定参数集不变．举个例子，我们考虑模型价格为 $h(x_t;\Theta)=bx_t$ 的情况．我们考虑以下参数集：$a=0.97$，$b=1.01$，$\lambda=0.1$，$\sigma=0.25$．在图 25-9a 和图 25-9b 中我们分别给出了隐藏状态和观察值的样本路径．在图 25-10 中我们展示了通过滤波技术得到的隐藏状态，并与原始隐藏状态对比．我们使用了 100 个样本，$m=100$．第二种情况是我们没有参数集，仅根据观察到的时间序列来估计．有很多种方法可以完成这一步．其中一种方法是从参数集的初始估计出发．已知初始估计之后，我们可以沿用上一个例子中的方法，预测下一步的观察值，计算真实观察值与预测值的距离，计算总距离，并将其作为似然函数的替代．下面我们的目标是找到合适的

参数集使该替代函数取最小值，该参数集即为最优参数集．有几个需要注意的地方：(i)已知参数集的新猜测值之后，我们必须执行滤波步骤；(ii)隐藏状态的时间序列是该过程的副产品．

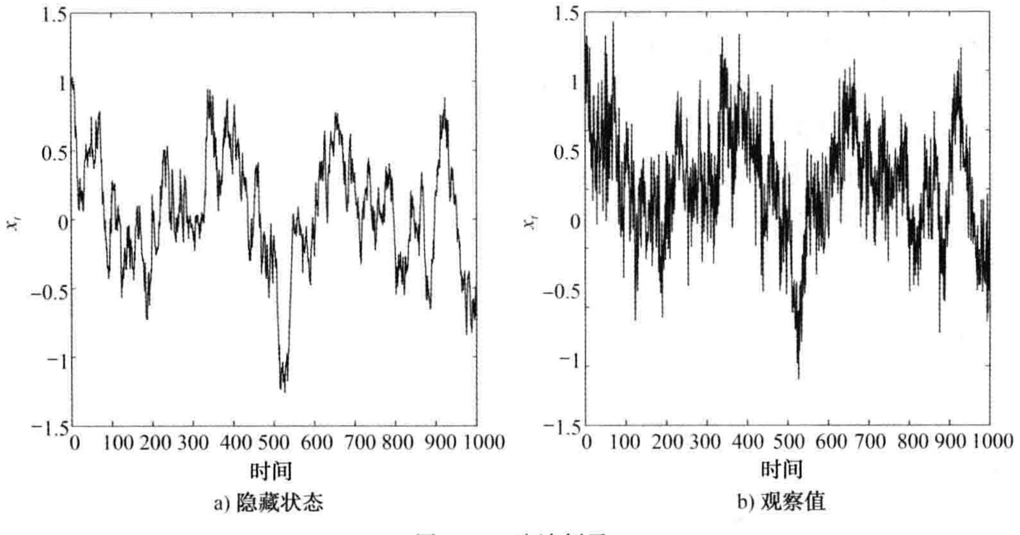


图 25-9 滤波例子

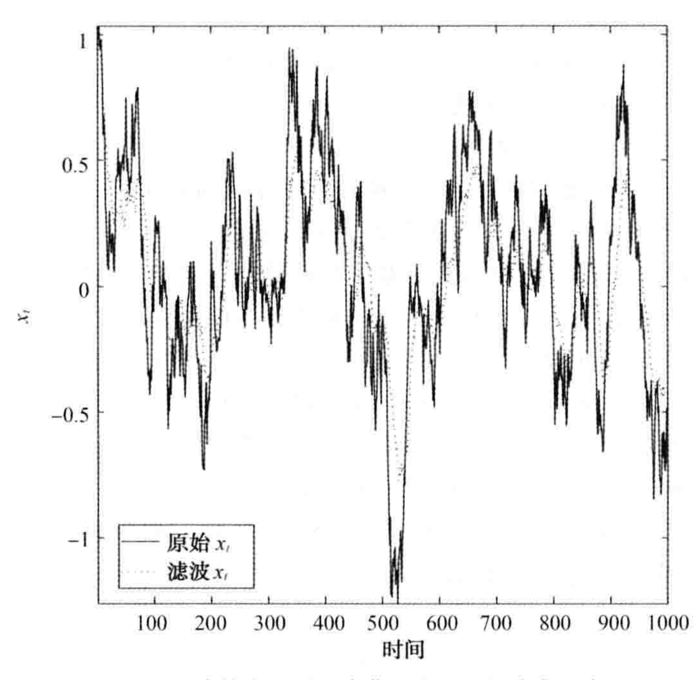


图 25-10 滤波技术得到的隐藏状态与原始隐藏状态对比

25.3.1 Kalman 滤波

430～431

Kalman 滤波是一种更加正式的估计方法，包括线性 Kalman 滤波和一些扩展版本，如扩展 Kalman 滤波、无损 Kalman 滤波、非线性情况下的粒子滤波.

假设状态变化服从以下线性方程

$$x_k = F_k x_{k-1} + w_k \tag{25.36}$$

其中 x_k 是 t_k 时刻的真实状态.

F_k 是状态转移矩阵，假设与时间有关.

w_k 是过程噪声，假设服从多变量正态分布，均值为 0，协方差为 Q_k，即 $w_k \sim \mathcal{N}(0, Q_k)$.

这里 x_k 是 $n\times1$ 的实值向量，F_k 为 $m\times n$ 的矩阵. 在 k 时刻，x_k 状态下的观察值或测量值 y_k 的演化服从以下测量方程：

$$y_k = H_k x_k + v_k \tag{25.37}$$

其中 H_k 是观察矩阵，v_k 是观察噪声，假设噪声为白噪声(高斯的)，均值为 0 协方差为 R_k，即

$$v_k \sim \mathcal{N}(0, R_k) \tag{25.38}$$

我们假设初始状态为 x_0，状态噪声向量为 $w_1, \cdots, w_k$，测量噪声向量为 $v_1, \cdots, v_k$，它们都是相互独立的. 在现实中，一般不能事先假定动态系统服从以上框架. 在这样的设定下，将 Kalman 滤波设计为能够在噪声环境下运行，能够使得该工具更为实用. 此外可以对 Kalman 滤波进行变化和拓展，得到更加成熟的模型.

我们用以下符号表示滤波的状态：

$\hat{x}_{k\mid k+1}$表示对 k 时刻状态的估计，基于截止到 $k-1$ 时刻的观察. x 的下标应该是 $k\mid k-1$

$\hat{x}_{k\mid k}$表示对 k 时刻状态的估计，基于截止到 k 时刻的观察.

$P_{k\mid k+1}$表示 k 时刻的误差协方差矩阵，基于截止到 $k-1$ 时刻的观察. 下标应该是 $k\mid k-1$

$P_{k\mid k}$表示 k 时刻的误差协方差矩阵，基于截止到 k 时刻的观察.

误差协方差矩阵能够衡量对状态估计的估计精度. 正规的滤波通过两个步骤完成：(a)时间更新或预测，(b)测量更新或滤波. 在预测环节中，我们利用来自过去时间步的状态估计得到当前时间步的状态估计. 在滤波环节中，我们使用当前时间步的测量值来优化该预测，从而得到新的当前时间步(理论上)更加准确的状态估计.

在我们执行这些步骤前，我们需要定义一些不变量. 如果模型是准确的，$\hat{x}_{0\mid 0}$和 $P_{0\mid 0}$准确反映初始状态的分布，那么以下不变量保持不变：(a)所有估计的误差均值为 0，即

$$E(x_k - \hat{x}_{k\mid k}) = 0 \tag{25.39}$$

$$E(x_k - \hat{x}_{k\mid k-1}) = 0 \tag{25.40}$$

以及(b)协方差矩阵准确反映估计的协方差，即

$$P_{k|k}=\mathrm{cov}(x_k-\hat{x}_{k|k}) \tag{25.41}$$

$$P_{k|k-1}=\mathrm{cov}(x_k-\hat{x}_{k|k-1}) \tag{25.42}$$

1. 时间更新

在 k 时刻，我们已知过去时刻状态的估计值，即 $\hat{x}_{k-1|k-1}$. 已知这些信息，我们只需将 $\hat{x}_{k-1|k-1}$带入方程(25.36)，不考虑噪声，即可得到 k 时刻状态的预测. 因此预测状态的估计值为

$$\boxed{\hat{x}_{k|k-1}=F_k\hat{x}_{k-1|k-1}} \tag{25.43}$$

已知预测状态的估计值，我们现在希望计算预测估计协方差矩阵，即 $P_{k|k-1}=\mathrm{cov}(x_k-x_{k|k-1})$. 432

$$\begin{aligned}
P_{k|k-1}&=\mathrm{cov}(x_k-\hat{x}_{k|k-1}) && P_{k|k-1}\text{ 的定义}\\
&=\mathrm{cov}(x_k-F_k\hat{x}_{k-1|k-1}) && \text{用(25.43)替换}\hat{x}_{k|k-1}\\
&=\mathrm{cov}(F_k\hat{x}_{k-1}+w_k-F_k\hat{x}_{k-1|k-1}) && \text{用(25.36)替换}x_k\\
&=\mathrm{cov}(F_k(x_{k-1}-\hat{x}_{k-1|k-1})+w_k) && \text{合并同类项}\\
&=\mathrm{cov}(F_k(x_{k-1}-\hat{x}_{k-1|k-1}))+\mathrm{cov}(w_k) && \text{假设两项独立}\\
&=F_k\mathrm{cov}((x_{k-1}-\hat{x}_{k-1|k-1}))F_k^{\mathrm{T}}+Q_k && \text{利用公式 }\mathrm{cov}(Ax)=A\mathrm{cov}(x)A^{\mathrm{T}}\\
&=F_kP_{k-1|k-1}F_k^{\mathrm{T}}+Q_k && P_{k-1|k-1}\text{ 的定义}
\end{aligned}$$

因此预测估计协方差为

$$\boxed{P_{k|k-1}=F_kP_{k-1|k-1}F_k^{\mathrm{T}}+Q_k} \tag{25.44}$$

2. 测量更新

已知预测状态估计 $x_{k|k-1}$，我们可以将其带入测量方程(25.37)，计算测量余量，即

$$\boxed{\hat{\delta}_k=y_kH_k\hat{x}_{k|k-1}} \tag{25.45}$$

令 S_k 为余量协方差，我们可以按以下步骤计算

$$\begin{aligned}
S_k&=\mathrm{cov}(\hat{\delta}_k) && S_k\text{ 的定义}\\
&=\mathrm{cov}(y_k-H_k\hat{x}_k\mid x_{k-1}) && \text{将(25.45)带入}\\
&=\mathrm{cov}(H_kx_k+v_k-H_k\hat{x}_{k|k-1}) && \text{将(25.37)带入}\\
&=\mathrm{cov}(H_k(x_k-\hat{x}_{k|k-1})+v_k) && \text{重组}\\
&=\mathrm{cov}(H_k(x_k-\hat{x}_{k|k-1}))+\mathrm{cov}(v_k) && \text{假设两项独立}\\
&=H_k\mathrm{cov}(x_k-\hat{x}_{k|k-1})H_k^{\mathrm{T}}+R_k && \text{利用公式 }\mathrm{cov}(Ax)=A\mathrm{cov}(x)A^{\mathrm{T}}\\
&=H_kP_{k|k-1}H_k^{\mathrm{T}}+R_k && P_{k|k-1}\text{ 的定义}
\end{aligned}$$

因此，余量的协方差矩阵为

$$\boxed{S_k=H_kP_{k|k-1}H_k^{\mathrm{T}}+R_k} \tag{25.46}$$

为了更新状态估计，可以简单地将余量 $\hat{\delta}_k$ 加到原先的估计值 $\hat{x}_{k|k-1}$，或者我们可以有更一般的形式：

$$\boxed{\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k \hat{\delta}_k \hat{x}_k} \tag{25.47}$$

其中 K_k 依然被称作为 Kalman. 到现在为止，我们并没有做过任何最优性的假设．现在我们考虑更新后的估计值协方差矩阵 $P_{k|k}$，也可以称为后验协方差矩阵．我们可以如下地计 433 算这个协方差

$$
\begin{aligned}
P_{k|k} &= \mathrm{cov}(x_k - \hat{x}_{k|k}) && P_{k|k} \text{ 的定义} \\
&= \mathrm{cov}(x_k - (\hat{x}_{k|k-1} + K_k \hat{\delta}_k)) && \text{将(25.47)代入} \\
&= \mathrm{cov}(x_k - (\hat{x}_{k|k-1} + K_k(y_k - H_k \hat{x}_{k|k-1}))) && \text{利用(25.45)替换 } \hat{\delta}_k \\
&= \mathrm{cov}(x_k - (\hat{x}_{k|k-1} + K_k(H_k x_k + v_k - H_k \hat{x}_{k|k-1}))) && \text{利用(25.37)替换 } y_k \\
&= \mathrm{cov}((I - K_k H_k)(x_k - \hat{x}_{k|k-1}) - K_k v_k) && \text{重新组合}
\end{aligned}
$$

由于假设测量误差与其他项不相关，我们可得

$$P_{k|k} = \mathrm{cov}((I - K_k H_k)(x_k - \hat{x}_{k|k-1})) + \mathrm{cov}(K_k v_k) \tag{25.48}$$

$$= (I - K_k H_k)\mathrm{cov}(x_k - \hat{x}_{k|k-1})(I - K_k H_k)^{\mathrm{T}} + K_k \mathrm{cov}(v_k) K_k^{\mathrm{T}} \tag{25.49}$$

利用 $P_{k|k-1}$ 和 R_k 的定义，更新的估计协方差为

$$\boxed{P_{k|k} = (I - K_k H_k) P_{k|k-1} (I - K_k H_k)^{\mathrm{T}} + K_k R_k K_k^{\mathrm{T}}} \tag{25.50}$$

该式对任意取值的 K_k 都成立．截至目前我们没有任何关于 K_k 的假设．为得到最优 Kalman 增益，我们可以进一步简化该表达式．

25.3.2 最优 Kalman 增益、含义及后验协方差矩阵

Kalman 滤波器需要使后验状态估计最小化，也就是说使均方误差估计量最小．后验状态估计的误差为

$$\epsilon_k \equiv x_k - \hat{x}_{k|k} \tag{25.51}$$

目标是使 ϵ_k 平方的期望值最小

$$E(|x_k - \hat{x}_{k|k}|^2) \tag{25.52}$$

这等价于使后验估计协方差矩阵 $P_{k|k}$ 的迹最小．$n \times n$ 方阵 A 的迹定义为

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} a_{ii} \tag{25.53}$$

即对角线上元素的和．将 $P_{k|k}$ 的项展开后，可得以下式子：

$$P_{k|k} = P_{k|k-1} - K_k H_k P_{k|k-1} - P_{k|k-1} H_k^{\mathrm{T}} K_k^{\mathrm{T}} + K_k (H_k P_{k|k-1} H_k^{\mathrm{T}} + R_k) K_k^{\mathrm{T}} \tag{25.54}$$

$$= P_{k|k-1} - K_k H_k P_{k|k-1} - P_{k|k-1} H_k^{\mathrm{T}} K_k^{\mathrm{T}} + K_k S_k K_k^{\mathrm{T}} \tag{25.55}$$

现在对 $P_{k|k}$ 的迹取一阶导，并令它等于 0，可以得到最优值 K_k：

$$\frac{\partial \mathrm{Tr}(P_{k|k})}{\partial K_k} = -2(H_k P_{k|k-1})^{\mathrm{T}} + 2K_k S_k = 0 \tag{25.56}$$

求解 K_k 可得

$$K_kS_k = (H_kP_{k|k-1})^{\mathrm{T}} = P_{k|k-1}H_k^{\mathrm{T}} \tag{25.57}$$

或

$$K_k = P_{k|k-1}H_k^{\mathrm{T}}K_k^{\mathrm{T}} \tag{25.58}$$

这是通过最小化均方误差估计量得到的 Kalman 增益，也是最优 Kalman 增益：

$$K_kS_kK_k^{\mathrm{T}} = P_{k|k-1}H_k^{\mathrm{T}}K_k^{\mathrm{T}} \tag{25.59}$$

代回到 $P_{k|k}$ 可得

$$P_{k|k} = P_{k|k-1} - K_kH_kP_{k|k-1} - P_{k|k-1}H_k^{\mathrm{T}}K_k^{\mathrm{T}} + K_kS_kK_k^{\mathrm{T}} \tag{25.60}$$

434

注意到最后两项可以抵消掉，可得

$$P_{k|k} = P_{k|k-1} - K_kH_kP_{k|k-1} \tag{25.61}$$

$$= (I - K_kH_k)P_{k|k-1} \tag{25.62}$$

我们可以基于线性回归来理解 Kalman 滤波器．在已知时间序列 $\{x_k\}$ 和 $\{y_k\}$ 的情况下，线性回归可得

$$y_k = \alpha + \beta x_k + \epsilon_k \tag{25.63}$$

α 为截距，β 为斜率，ϵ_k 为残差．在线性回归下我们有

$$\beta = P_{k|k-1}H_k^{\mathrm{T}}S_k^{-1} \tag{25.64}$$

该式即为 Kalman 增益的表达式．

在 Kalman 滤波器中，各时间步的对数似然函数为 $\log(z_t|z_{1:t-1})$．这个函数可以通过计算一个均值为 0，协方差为 S_k 的多变量高斯分布的对数概率分布函数得到，其中 S_k 的计算基于 $\hat{\delta}_k$ 的值．那么，对数似然函数为

$$\log p(y_t|y_{1:t-1}) = \log\left(\frac{1}{(2\pi)^{d/2}\,|S_k|^{1/2}}\exp\left(-\frac{1}{2}(\hat{\delta}_k - 0)^{\mathrm{T}}S_k^{-1}(\hat{\delta}_k - 0)\right)\right) \tag{25.65}$$

$$= -\frac{2}{d}\log(2\pi) - \frac{1}{2}\log|S_k| - \frac{1}{2}(\hat{\delta}_k - 0)^{\mathrm{T}}S_k^{-1}(\hat{\delta}_k - 0) \tag{25.66}$$

$$= -\frac{2}{d}\log(2\pi) - \frac{1}{2}\log|S_k| - \frac{1}{2}(z_k - H_k\hat{x}_{k|k-1})^{\mathrm{T}}S_k^{-1}(z_k - H_k\hat{x}_{k|k-1}) \tag{25.67}$$

记该对数似然函数为 $\ln(l_k)$，那么整个时间序列的对数似然函数为 $\ln(L_{1:T}) = \sum_{k=1}^{T}\ln(l_k)$．已知对数似然函数后，我们可以使用优化方法(如 Nelder-Mead 法)最大化(或最小化负的)对数似然函数，从而求出参数估计量．必须注意 Kalman 增益不会成为参数估计；它只进入滤波步骤.

25.4　习题

1. 已知 Kalman 滤波器的对数似然方程如(25.67)所示，利用它估计本章的例子中线性模型的参数．

2. 在对同一数据集的校准过程中，为观察目标函数的选择对所得参数集的影响，我们使用不同的目标函

数，重新完成本章的例子．使用表 25-1 和表 25-2 中的期权数据，针对以下不同的目标函数，校准不同期限虚值期权的方差伽马参数.

(a) $\sum_{i=1}^{n}(C_i^{\Theta}-C_i)^2$

(b) $\sum_{i=1}^{n}|C_i^{\Theta}-C_i|$

(c) $\sum_{i=1}^{n}\left(\frac{C_i^{\Theta}-C_i}{C_i}\right)^2$

(d). $\sum_{i=1}^{n}\left|\frac{C_i^{\Theta}-C_i}{C_i}\right|^2$

(e) $\sum_{i=1}^{n}w_i(C_i^{\Theta}-C_i)^2$，其中 $w_i=\frac{1}{(C_i^{\text{aks}}-C_i^{\text{bid}})^2}$

(f) $\sum_{i=1}^{n}w_i(C_i^{\Theta}-C_i)^2$，其中 w_i 定义同上

3. 利用表 25-1 和表 25-2 中的数据以及方程(25.12)，构建局部波动率曲面．在构建之前，先设定统一的网格，然后使用双三次样条向内/向外插值，为这些格点找到取值，然后使用方程(25.12)估计每个格点的局部波动率．可以查找一些估算偏导数的方法．

435

参考文献

Andreasen, Jesper, 1998. Implied Modelling, Stable Implementation, Hedging, and Duality. Manuscript, University of Aarhus.

Arvanitis, A., Gregory, J., Laurent, J.P., 1999. Building models for credit spreads. Journal of Derivatives,.

Bertoin, Jean, 1996. Lévy Processes. Cambridge University Press, Cambridge. Tracts in Mathematics 121.

Carr, Peter, Madan, Dilip B., 1999. Option valuation using the fast Fourier transform. Journal of Computational Finance 2 (4), 61–73.

Carr, Peter, Madan, Dilip B., 2009. Saddlepoint methods for option pricing. Journal of Computational Finance 13 (1) Fall.

Carr, Peter, Geman, Hélyette, Madan, Dilip B., Yor, Marc, 2002. The fine structure of asset returns: an empirical investigation. Journal of Business 75 (2), 305–332.

Carr, Peter, Geman, Hélyette, Madan, Dilip B., Yor, Marc, 2003. Stochastic volatility for Lévy processes. Mathematical Finance 13 (3), 345–382.

Cont, Rama, 2010. Model Calibration. In: Cont, Rama (Ed.). Encyclopedia of Quantitative Finance, vol. 3. John Wiley and Sons Ltd., Southern Gate, Chichester, West Sussex, England, pp. 1210–1218.

Cont, Rama, Tankov, Peter, 2003. Financial Modelling with Jump Processes. Chapman & Hall/CRC Financial Mathematics Series.

Cooley, James W., Tukey, John W., 1965. An algorithm for the machine calculation of complex Fourier series. Mathematics of Computation 19 (90), 297–301.

Derman, Emanuel, Kani, Iraj, 1994. The volatility smile and its implied tree. Risk 7 (2), 32–39.

Dupire, Bruno, 1994. Pricing with a smile. Risk 7 (1), 18–20.

Gregory, Jon, Laurent, Jean-Paul, 2005. Basket Default Swaps, CDOs and Factor Copulas. Review of Risk 7 (4), 103–122.

Grimmett, Geoffrey R., Stirzaker, David R. , 1992. Probability and Random Processes, second ed. Oxford University Press Inc., New York, USA.

Hirsa, Ali, 2012. Computational Methods in Finance. Chapman & Hall/CRC Financial Mathematics Series.

Hatgioannides, John, Petropoulos, George, 2007. On credit spreads, credit spread options and implied probabilities of default. Faculty of Finance, Cass Business School, UK and Eurobank EFG, Athens, Greece (January).

Hirsa, Ali, Courtadon, Georges, Madan, Dilip B., March 2001. Local Volatility Reengineers Semi-martingle Models. Morgan Stanley.

Hirsa, Ali, Courtadon, Georges, Madan, Dilip B., 2003. The effect of model risk on the valuation of barrier options. Journal of Risk Finance 4 (2).

Jacod, Jean, Shiryaev, Albert N. , 1987. Limit Theorems for Stochastic Processes. Springer-Verlag, Berlin, New York.

Jarrow, Robert A., Thurnbull, Stuart M., 1995. Pricing of options on financial securities subject to credit risk. Journal of Finance, 53–85.

Jarrow, Robert A., Turnbull, Stuart M., 1995. Pricing derivatives on financial securities subject to credit risk. Journal of Finance 50 (1), 53–85.

Longstaff, Francis A., Schwartz, Eduardo S., 1992. Interest rate volatility and the term structure: a two-factor general equilibrium model. Journal of Finance 47, 1259–1282.

Longstaff, Francis A., Schwartz, Eduardo S., 1995. A simple approach to valuing risky fixed and floating rate debt. Journal of Finance 50 (3), 789–819.

Madan, Dilip B., Carr, Peter, Chang, Eric C., 1998. The variance gamma process and option pricing. European Finance Review 2, 79–105.

Madan, Dilip B., Unal, Haluk, 2000. A two-factor hazard rate model for pricing risky debt and the term structure of credit spreads. Journal of Financial and Quantitative Analysis 35 (1), 43–65.

Merton, Robert C., 1976. Option pricing when underlying stock returns are discontinuous. Journal of Financial Economics 3, 125–144.

Øksendal, Bernt , 2000. Stochastic Differential Equations. An Introduction with Applications, fifth ed. Springer-Verlag, New York.

Papapantoleon, Antonis, 2008. An Introduction to Lévy Processes with Applications in Finance. Lecture notes taught at the University of Piraeus, University of Leipzig, and at the Technical University of Athens.

Protter, Philip , 1990. Stochastic Integration and Differential Equations. Springer-Verlag, New York.

Sato, Ken-iti, 2001. Basic results on Lévy processes. In: Barndorff-Nielsen, Ole E., Mikosch, Thomas, Resnick, Sidney I. (Eds.), Lévy Processes: theory and Applications. Birkhauser, Boston, MA, USA, pp. 3–37.

索　引

索引中的页码为英文原书页码，与书中页边标注的页码一致。

C

D

G

H

L

M

N

O

P

S

T

U

V

W

Y

Z

金融衍生工具数学导论（原书第3版）

在修订Neftci的著作第3版的工作中，Ali Hirsa取得了杰出的成果。较上一版本，本书增加了许多新的章节，如信用衍生品（第23章）、跳跃过程、偏积分-微分方程等。此外还有数值计算方法方面的新内容，如傅里叶变换（第22章）、校准技术（第25章），并增加了相应的例子和练习。总之，本书比上一个版本在各章上都有明显的提升。作为一本内容精、受众广的金融数学导论，本书是一本罕有的佳作。非常赞！

——**Jean-Pierre Fouque，加利福尼亚大学**

一位受人敬仰的学者为一部经典著作推出了内容更前沿的版本，无论对从业者还是学术圈，这都是件不折不扣的好事！

——**Lars Tyge Nielsen，哥伦比亚大学**

介绍数学、概率、随机微积分的著作数不胜数，然而鲜有聚焦于金融衍生品的对冲和定价这一课题的。多年来我一直使用本书的上一版本作为教材，以后我会用这个新版本。毫无疑问，本书将继续为金融从业者提供可贵的参考。

——**Robert L. Kimmel，新加坡国立大学**

本书介绍了一系列如何用最可行的数学方法对金融衍生品定价的问题。作者在介绍与金融市场相关的基础数学概念时非常注重启发性，这能有效地促使读者在衍生品定价的学习中打下良好基础。

——**Seppo Pynnonen，瓦萨大学**

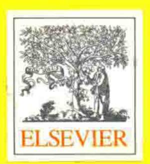
对于缺乏数学或实际应用背景的学生或从业者来说，本书最可贵之处在于它在对概念娓娓道来的同时，还注重读者的启发式学习，使其更容易明白金融问题的数学解决方法。相比第2版，本书增加了信用衍生品和PDE等内容，紧跟时代的节奏。

——**Mishael Milaković，班贝克大学**

作者简介

艾利·赫萨（Ali Hirsa）是应用数学博士，哥伦比亚大学与纽约大学兼职教授，先后工作于摩根士丹利、美银证券、保诚证券等多家大型金融机构。

萨利赫 N. 内夫特奇（Salih N.Neftci）在明尼苏达大学获得博士学位，生前先后在乔治华盛顿大学、波士顿大学、日内瓦国际研究机构、纽约城市大学研究生院和英国瑞丁大学isma中心任教。除了教学之外，他还参与金融机构有关定价和风险管理方面的研究工作。不仅在学术刊物上发表了多篇影响广泛的学术论文，而且一直是数家大银行和国际机构的咨询顾问。



An Introduction to the Mathematics of Financial Derivatives

Third Edition

本书译自原版*An Introduction to the Mathematics of Financial Derivatives, Third Edition*并由Elsevier授权出版







定价：99.00元

投稿热线：(010) 88379604
客服热线：(010) 88378991 88361066
购书热线：(010) 68326294 88379649 68995259

华章网站：www.hzbook.com
网上购书：www.china-pub.com
数字阅读：www.hzmedia.com.cn

封面设计·杨宇梅

[General Information]
书名=金融衍生工具数学导论 原书第3版
作者=（美）艾利?赫萨，（美）萨利赫 N.内夫特奇著
页数=442
SS号=14044062
DX号=
出版日期=2016.08
出版社=机械工业出版社